

This report was prepared by Westat using federal funding provided by the Bureau of Justice Statistics.

Document Title: Methodological Research to Support the National Crime Victimization Survey:
Self-Report Data on Rape and Sexual Assault – Pilot Test

Authors:	David Cantor, Westat Darby Miller Steiger, Westat Reanne Townsend, Westat John Y. Hartge, Westat Robert E. Fay, Westat Antonia Warren, Westat Leanne L. Heaton, Westat	Suzanne Kaasa, Westat Aaron Maitland, Westat Hanyu Sun, Westat Greg Norman, Westat Michael Jones, Westat Shannan Catalano, BJS (Former) Allen J. Beck, BJS
----------	--	--

Document No.: NCJ 256011

Publication Date: January 2021

Award No.: This project was supported by award number 2011-NV-CX-K074.

Abstract:

This paper describes the methodology, key findings, recommendations, and limitations of the Rape and Sexual Assault Pilot Test. The Bureau of Justice Statistics initiated the study to develop recommendations on the best methods to collect data on rape and sexual assault within the National Crime Victimization Survey program, as data collections using different methodologies and definitions have generated competing estimates on the level of sexual victimization in the United States over the past two decades.

Disclaimer

The Bureau of Justice Statistics funded this third-party report. It is not a BJS report and does not release official government statistics. The report is released to help inform interested parties of the research or analysis contained within and to encourage discussion. The results of the pilot test are not intended to provide official estimates of the prevalence or incidence of rape or sexual assault in the five metropolitan areas selected for the study. BJS has performed a limited review of the report to ensure the general accuracy of information and adherence to confidentiality and disclosure standards. Any statistics included in this report are not official BJS statistics unless they have been previously published in a BJS report. Any analysis, conclusions, or opinions expressed herein are those of the authors and do not necessarily represent the views, opinions, or policies of the Bureau of Justice Statistics or the U.S. Department of Justice. See *Bureau of Justice Statistics Assessment of the Rape and Sexual Assault Pilot Test. (2021)*. BJS website.

This page intentionally left blank.

Methodological Research to Support the National Crime Victimization Survey: Self-Report Data on Rape and Sexual Assault – Pilot Test

August 2018

Authors

Westat

David Cantor
Darby Miller Steiger
Reanne Townsend
John Y. Hartge
Robert E. Fay
Antonia Warren

Leanne L. Heaton
Suzanne Kaasa
Aaron Maitland
Hanyu Sun
Greg Norman
Michael Jones

BJS Project Staff

Shannan Catalano (Former)
Allen J. Beck



Prepared for
Bureau of Justice Statistics
810 Seventh Street, NW
Washington, DC 20531

Prepared by
Westat
An Employee-Owned Research Corporation®
1600 Research Boulevard
Rockville, Maryland 20850-3129
(301) 251-1500

Table of Contents

<u>Chapter</u>		<u>Page</u>
	Disclaimer and Acknowledgments	xxix
	Executive Summary.....	xxxii
	Key Findings – Data Quality.....	xxxv
	Key Findings – Comparison to the NCVS	xlvi
	Recommendations	xlviii
	Limitations and Additional Research.....	lvii
1	Background and Purpose	1
1.1	Critical Design Components for Rape and Sexual Assault Surveys.....	3
1.1.1	Mode of Interview and Maximizing Privacy and Confidentiality.....	4
1.1.2	Minimize Coverage Error and Nonresponse Bias.....	5
1.1.3	Minimize Measurement Error Associated with the Questionnaire	7
1.1.4	Minimize Emotional Trauma.....	12
1.2	Overview of this Report	14
2	Development of the Surveys.....	17
2.1	Overview of Sample Designs	17
2.1.1	General Population Sample.....	17
2.1.2	Volunteer Sample	19
2.1.3	Service Provider Sample	20
2.2	Overview of survey procedures	20
2.2.1	In-Person Survey: General Population Sample.....	20
2.2.2	Telephone Survey: General Population Sample.....	21
2.2.3	In-Person Survey: Volunteer and Service Provider Samples	21
2.2.4	Telephone Survey: Volunteer and Service Provider Samples	22
2.2.5	Re-Interview	22

Contents (continued)

<u>Chapter</u>		<u>Page</u>
2.3	Surveys to Measure Rape and Sexual Assault.....	23
2.3.1	Defining and Screening Incidents as Rape and Sexual Assault.....	25
2.3.2	Detailed Incident Form	28
2.4	Structure of the Interviews	31
2.5	Tests of Procedures and Surveys.....	33
2.5.1	Cognitive Testing.....	33
2.5.2	Feasibility Study	35
3	Sampling and Weighting.....	38
3.1	Initial Sample Designs and Goals.....	39
3.2	Further Details on the ACASI Design.....	46
3.3	Further Details on the CATI Design.....	50
3.4	Modifications to the ACASI Design During the Field Period.....	53
3.5	ACASI Weighting	56
3.6	CATI Weighting.....	66
4	Data Collection	77
4.1	In-Person – General Population Sample.....	77
4.1.1	Household Screening – Roster Mailing.....	78
4.1.2	Household Screening – In-Person Visits	81
4.1.3	Respondent Interviews	82
4.1.4	Respondent Re-Interviews	82
4.1.5	Respondent Incentives.....	83
4.1.6	University Sample.....	83
4.2	Telephone – General Population Sample.....	85
4.2.1	Screening.....	85
4.2.2	Respondent Interviews	86
4.2.3	Respondent Incentives.....	87

Contents (continued)

<u>Chapter</u>		<u>Page</u>
4.3	Volunteer Sample.....	87
4.3.1	Recruitment	87
4.3.2	Interviewing the Volunteer Sample	89
4.4	Service Provider Sample	90
4.4.1	Service Provider (SP) Respondent Recruitment	91
4.4.2	Administration of the ACASI Interviews	92
4.4.3	Administration of the Telephone Interviews	92
4.5	Field Interviewer Training and Monitoring	93
4.5.1	Interviewer Training.....	93
4.5.2	Field Interviewing Organization.....	94
4.6	Telephone Interviewer Training/Monitoring.....	96
4.6.1	Interviewer Training.....	96
4.6.2	Telephone Interviewing Organization.....	97
4.7	Field Problems – Difficulty Reaching Respondents.....	98
4.7.1	GP Sample	98
4.7.2	VO and SP Sample.....	99
4.8	Telephone Problems – Difficulty Reaching Respondents	100
4.8.1	RDD Sample	100
4.8.2	VO Sample	101
4.8.3	SP Sample	102
5	Response Rates	103
5.1	Overall Unweighted Response Rates	103
5.2	Types of Nonresponse	111
5.3	Weighted Response Rates.....	114

Contents (continued)

<u>Chapter</u>		<u>Page</u>
6	Timing and Characteristics of Respondents.....	118
	6.1 Time to Complete the Survey	118
	6.2 Demographics of Respondents.....	121
	6.2.1 Race and Hispanic Origin.....	122
	6.2.2 Age.....	123
	6.2.3 Education.....	126
	6.2.4 Income	127
	6.2.5 Marital Status and Relationship Status	127
	6.2.6 Employment Status and Military Status	129
	6.2.7 Home Ownership and Length of Stay	129
	6.2.8 Public Behaviors	131
7	Prevalence and Incidence	134
	7.1 Incident Classification Algorithm.....	134
	7.1.1 Behavior	134
	7.1.2 Tactic	135
	7.1.3 List of all Possible Classification Codes	137
	7.1.4 Narrative Examples of Classification Categories.....	140
	7.1.5 Finalizing the Classification Scheme.....	142
	7.2 Sexual Victimization Rates.....	148
	7.2.1 Prevalence and Incidence of Sexual Victimization in the Past 12 Months	148
	7.2.2 Comparison to the NCVS and NISVS.....	157
	7.2.3 Summary of Results on Prevalence and Incidence	170
8	Characteristics of Victims and Incidents.....	174
	8.1 Characteristics of Victims	174
	8.2 Circumstances of the Incident	179
	8.2.1 Time of Day of Incident.....	180
	8.2.2 Location of Incident.....	182
	8.2.3 Use of Verbal Pressure and Coercive Tactics In Incident	183

Contents (continued)

<u>Chapter</u>		<u>Page</u>
8.2.4	Use of Force	186
8.2.5	Injury From Incident.....	188
8.2.6	Offender Characteristics.....	190
8.2.7	Offender Use of Alcohol and/or Drugs During the Incident.....	194
8.2.8	Respondent Use of Alcohol and/or Drugs During the Incident.....	196
8.2.9	Efforts to Try to Stop the Incident.....	205
8.2.10	Did the Offender Immediately Stop When You Said You Didn't Want it to Happen?	208
8.2.11	Level of Distress	209
8.2.12	Emotions Felt as Result of the Incident	211
8.2.13	Perceptions of the Incident as a Crime at the Time of the Incident	214
8.2.14	Police Involvement.....	218
8.2.15	Involvement of Others	220
8.3	Comparing Circumstances for the RSA Pilot Test and NCVS.....	228
8.3.1	Time of Day of Incident.....	229
8.3.2	Location of Incident.....	230
8.3.3	Use of a Weapon	231
8.3.4	Injury from Incident.....	232
8.3.5	Number of Offenders	238
8.3.6	Sex of Single Offender.....	239
8.3.7	Relationship to Offender.....	240
8.3.8	Offender Use of Alcohol or Drugs.....	242
8.3.9	Respondent Emotions After Incident	243
8.3.10	Police Informed About Incident.....	244
8.3.11	Involvement of Others	245
8.4	Summary and Discussion.....	247
8.4.1	Who are the Victims?	247
8.4.2	What are the Circumstances of Victimization for the RSA Pilot Test?	248
8.4.3	How do the Incidents on RSA Pilot Test Compare to Those on the NCVS?	249

Contents (continued)

<u>Chapter</u>		<u>Page</u>
9	Assessment of Two-Stage Design for Measuring RSA.....	252
9.1	Sexual Victimization Screener.....	252
9.1.1	Incident Eligibility and Reports of Duplicates	252
9.1.2	Incident Date Distribution and Telescoping.....	259
9.1.3	Missing Data.....	265
9.2	Detailed Incident Form Results.....	268
9.2.1	DIF Eligibility and Completion.....	268
9.2.2	Missing Data.....	273
9.3	Comparing BSQ, DIF and Narratives	274
9.3.1	Comparison of Classifications for Completed Rape	279
9.3.2	Comparison of Classifications for Attempted and Threatened Penetration.....	292
10	Nonresponse Bias and Related Concerns	304
10.1	Effects of Differential Response Rates.....	305
10.2	Neighborhood Characteristics of Respondents and Nonrespondents.....	310
10.3	Geographic Precision of the CATI Sample	316
10.4	Level of Effort Analysis.....	322
11.	Analysis of Vignettes, Event History Calendar, and Interviewer Behavior	328
11.1	Analysis of Vignettes	328
11.1.1	Alcohol Vignettes	329
11.1.2	Non-Physical Pressure Vignettes	342

Contents (continued)

<u>Chapter</u>		<u>Page</u>
11.2	Event History Calendar.....	349
11.2.1	Design and Administration	349
11.2.2	Events Reported on the Calendar.....	350
11.2.3	Respondents' Use of the EHC During the Interview	351
11.3	Interviewer Performance	357
11.3.1	Behavior Coding.....	357
11.3.2	CATI Interviewer Effects	366
12	Data Reliability.....	371
12.1	Methodology.....	371
12.1.1	Sampling and Response Rates	372
12.1.2	Primary Measure of Reliability.....	373
12.1.3	Reinterview Weights and Statistical Testing.....	374
12.2	Results for the Screener	375
12.3	Reliability of the Detailed Incident Form.....	381
12.3.1	How Often is the Same Incident Reported at Both Interviews?	381
12.3.2	How Reliable are Measures of Behaviors and Tactics on the DIF?.....	383
12.4	Summary and Discussion.....	409
13	Cost Effectiveness: Comparisons of ACASI and CATI Designs	411
13.1	Relative Cost Analysis	411
13.2	Relative Cost by RDD Sample Type.....	414
14	Respondent Reactions and Interviewing Environment.....	417
14.1	Respondent Reactions to the Survey	417
14.1.1	Emotional Reactions to the Survey.....	418
14.1.2	Comparisons to Other Studies	419

Contents (continued)

<u>Chapter</u>		<u>Page</u>
	14.1.3 Victimization	420
	14.1.4 Survey Ease of Understanding.....	426
14.2	Interviewer Observations	428
	14.2.1 Interviewer Ratings of Distress	429
	14.2.2 Signs of Distress and Cooperation.....	432
	14.2.3 Someone Else Present During the Interview	432
	14.2.4 ACASI Survey Setting and Verbal Interaction with Interviewer.....	437
	14.2.5 Discussion.....	447
15	Conclusions and Recommendations.....	449
	15.1 Summary of Results	449
	15.1.1 Mode and Privacy of the Interview.....	449
	15.1.2 Response and Coverage Rates	451
	15.1.3 Missing Data.....	452
	15.1.4 Measurement Error	453
	15.1.5 Minimize Emotional Trauma.....	458
	15.2 Comparisons with the NCVS.....	459
	15.2.1 How do RSA Pilot Test Estimates Differ from those of NCVS?	459
	15.2.2 Characteristics and Correlates of Victimization.....	461
	15.3 Recommendations	462
	15.3.1 Recommendations for Collection of Rape and Sexual Assault Data Within the NCVS Program Design.....	462
	15.3.2 Recommended Methods and Design Features of a Separate Survey	466
	15.3.3 Length and Structure of a Separate Survey.....	471
	15.3.4 Measurement of Key Concepts	473
	15.4 Summary and Limitations.....	480
	References.....	484

Contents (continued)

Tables		Page
E-1	Definition of false negative and false positive errors	xl
E-2	Incidence rate per 1,000 of rape and sexual assault comparing NCVS to RSA Pilot Test for females ages 18–49 in the 5 CBSAs	xliv
E-3	Percent of rapes and sexual assaults reported from BSQ items targeting different behaviors and tactics	xlvii
2-1	Questions used to screen for rape and sexual assault on the ACASI instrument.....	27
2-2	Topics covered on the detailed incident form.....	29
2-3	Content of interview by mode of data collection	33
2-4	Completion rates for feasibility study household roster, main interview and re-interview by sample type and interview mode, 2014	36
2-5	Approximate timings by section of the main interview by interview mode.....	37
3-1	Incidence rate of rape per 1,000 females and standard errors by age group, National Crime Victimization Survey, 2005–2010	41
3-2	Design assumptions for prevalence rates per 1,000 of rape and sexual assault over a 12-month period by interview mode for females ages 18–49 in the general population, 2014–2015	46
3-3	Number of exchanges, working banks, and telephone numbers in the landline and cell phone sampling frames by core-based statistical area, 2014–2015	52
3-4	Number of records selected and dialed in the landline and cell phone sampling frames by core-based statistical area, 2014–2015.....	53
3-5	Assignment status of wave 1 segments, included and excluded from the final segment sample	55

Contents (continued)

Tables		Page
3-6	Status of the wave 1 segments initially withheld in November 2014, by field region.....	55
3-7	Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015.....	60
3-8	Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015.....	63
3-9	Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015.....	71
3-10	Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015.....	74
4-1	Data collection periods	77
4-2	Completed main interviews by month	79
4-3	Outcome of ABS sample roster mailing by disposition.....	81
4-4	Participating colleges and universities.....	83
4-5	University sample results	84
4-6	Volunteer sample allocation by mode	88
4-7	Volunteer sample results.....	89

Contents (continued)

Tables		Page
4-8	Service provider sample productivity	92
4-9	Field staff organization	94
5-1	Calculation of unweighted response rates by core-based statistical area for the ACASI general population sample	106
5-2	Calculation of unweighted response rates by core-based statistical area (CBSA) for the CATI landline sample	109
5-3	Calculation of unweighted response rates by CBSA for the CATI cell sample and the combined overall rate for the CATI sample	110
5-4	Nonresponding households or persons to the screener by reason.....	112
5-5	Extended interview nonresponse by reason.....	115
5-6	Comparison of unweighted and weighted response rates by CBSA	117
6-1	Average time to complete in minutes by survey section, sample type, and mode, 2014–2015	119
6-2	Average time to complete extended interview by number of detailed incident forms, sample type, and mode, 2014–2015	121
6-3	Respondent demographics by mode of interview for females ages 18-49 in the general population, 2014–2015	123
6-4	Respondent demographics by mode of interview for females ages 18-29 in the volunteer sample, 2014–2015	125
6-5	Respondent romantic relationship status by mode of interview for females ages 18-49 in the general population, 2014–2015	128
6-6	Respondent romantic relationship status by mode of interview for females ages 18-29 in the volunteer sample, 2014–2015.....	129

Contents (continued)

Tables

<u>Tables</u>	<u>Page</u>
6-7 Respondent military status, employment, and living situation by mode of interview for females ages 18-49 in the general population, 2014–2015	130
6-8 Respondent military status, employment, and living situation by mode of interview for females ages 18-29 in the volunteer sample, 2014–2015	131
6-9 Respondent frequency of public behavior by mode of interview for females ages 18-49 in the general population, 2014–2015	132
6-10 Respondent frequency of public behavior by mode of interview for females ages 18-29 in the volunteer sample, 2014–2015.....	133
7-1 Behaviors identified by detailed incident form.....	135
7-2 Tactics identified by detailed incident form.....	137
7-3 Incident classification categories	138
7-4 Examples of narratives for key DIF classification categories	141
7-5 Examples of incidents involving attempted penetration in which classification was revisited based on narrative.....	144
7-6 Examples of incidents involving unwanted sexual contact in which classification was revisited based on narrative.....	145
7-7 Examples of incidents when the person immediately stopped in which classification was revisited based on narrative.....	147
7-8 Prevalence rate of unwanted sexual contact, by type of crime classification and mode of interview for females ages 18-49 in the general population, 2014–2015.....	149
7-9 Prevalence rate of rape and sexual assault by type of crime classification and mode of interview for females ages 18-49 in the general population, 2014–2015.....	150

Contents (continued)

Tables

	<u>Page</u>
7-10 Incidence rate of unwanted sexual contact per 1,000 females age 18-49, by type of crime classification and mode of interview in the general population, 2014–2015	153
7-11 Prevalence rate of unwanted sexual contact, by type of crime classification and mode of interview for females ages 18-29 in the volunteer sample, 2014–2015	155
7-12 Incidence rate per 1,000 of rape and sexual assault by study for females ages 18-49 in the general population.....	158
7-13 Comparison of prevalence rates from RSA Pilot Test to NCVS rates from 2011–2014 by type of victimization, study, and mode for females age 18-49 in the general population	159
7-14 Incidence rates of rape and sexual assault per 1,000 females age 18-49 by victimization screener item generating the report and whether respondent believed it was a crime, in the general population, 2014-2015	164
7-15 Comparison of prevalence rates from RSA Pilot Test to National Intimate Partner and Sexual Violence Survey (NISVS) rates, by type of victimization, reference period, mode, and classification method for females ages 18-49 in the general population.....	166
8-1 Significant predictors of rape and sexual assault rates, by characteristic, mode, and interaction	175
8-2 Odds ratios for significant predictors of rape victimization.....	178
8-3 Odds ratios for significant predictors of sexual assault victimization	179
8-4 Time of day incident occurred, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	181
8-5 Activity at time of incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	182

Contents (continued)

Tables		Page
8-6	Use of verbal pressure and coercive tactics, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	185
8-7	Force tactics, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	187
8-8	Physical injuries, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	188
8-9	Type of physical injuries, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	189
8-10	Medical care among those injured, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	190
8-11	Number of offenders, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	191
8-12	Sex of offender, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	192
8-13	The offender’s relationship to the victim by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	193
8-14	Offender use of alcohol or drugs leading up to incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	195
8-14a	Questionnaire items used to determine respondent level of substance use.....	197
8-15	Extent of substance use, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	198

Contents (continued)

Tables		Page
8-15a	Extent of substance use involved in incident by type of incident, level of intoxication and mode for females age 18-49 in the general population, 2014–2015.....	200
8-15b	Incidents of rape and sexual assault when the victim was unable to consent by classification and mode for females age 18-49 in the general population.....	202
8-15c	Responses to questions on intoxication for incidents when victim was unable to consent for females age 18-49 in the general population, 2014-2015	204
8-16	Respondent actions to try to stop incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	206
8-17	Offender immediately stopped when asked, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	208
8-18	Level of distress in days following incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	209
8-19	Level of distress at the time of interview, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	210
8-20	Emotions felt as result of the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	212
8-21	Respondent considered incident to be a crime at the time, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	214
8-22	Reasons why respondent did not consider incident to be a crime at the time, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	216

Contents (continued)

Tables

	<u>Page</u>
8-23 Respondent now considers incident to be a crime, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	217
8-24 Whether police were informed about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	218
8-25 Reasons why police were not informed about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	219
8-26 Whether respondent talked to a mental health professional, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	221
8-27 Whether respondent talked to a crisis hotline operator, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	222
8-28 Respondent told others about the incident, other than police, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	223
8-29 Whom respondent told about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	224
8-30 Whether respondent told any organization about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	225
8-31 What organization respondent told about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	226
8-32 Whether respondent received any help from an agency that deals with victims of crime, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	227

Contents (continued)

Tables

	<u>Page</u>
8-33 Type of help received, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	228
8-34 Time of day incident occurred, by study and mode of interview for females ages 18-49 in the general population	230
8-35 Location at time of incident, by study and mode of interview for females ages 18-49 in the general population.....	231
8-36 Use of a weapon, by study and mode of interview for females ages 18-49 in the general population	232
8-37 Physical injuries, by mode of interview for females ages 18-49 in the general population.....	234
8-38 Type of physical injuries, by mode of interview for females ages 18-49 in the general population.....	235
8-39 Medical care, by mode of interview for females ages 18-49 in the general population	236
8-40 Location of medical care, by mode of interview for females ages 18-49 in the general population.....	237
8-41 Whether respondent stayed overnight in the hospital, by mode of interview for females ages 18-49 in the general population	238
8-42 Number of offenders, by study and mode of interview for females ages 18-49 in the general population.....	239
8-43 Sex of single offender, by study and mode of interview for females ages 18-49 in the general population.....	240
8-44 The offenders relationship to the victim by study and mode of interview for females ages 18-49 in the general population	241
8-45 Offender's use of alcohol or drugs in the hours leading up to incident, by study and mode of interview for females ages 18-49 in the general population	243

Contents (continued)

Tables		Page
8-46	Emotions felt as result of the incident, by mode of interview for females ages 18-49 in the general population.....	244
8-47	Police notification, by study and mode of interview for females ages 18-49 in the general population	245
8-48	Whether respondent received any help from agency that deals with victims of crime by mode of interview for females ages 18-49 in the general population.....	246
9-1	Proportion of unique in scope incidents reported in victimization screener based on screener items, by mode of interview for females ages 18-49 in the general population, 2014–2015	254
9-2	Proportion of unique in scope incidents reported in victimization screener based on screener items, by mode of interview for females ages 18-49 in the volunteer sample, 2014–2015	257
9-3	Percent of respondents saying yes to each screener item who indicated the incident happened more than four times in the past 12 months by screener item and mode of interview for females ages 18-49 in the general population, 2014–2015.....	267
9-4	Number of DIF incidents completed by number of incidents eligible for a DIF and mode of interview for females ages 18-49 in the general population, 2014–2015.....	270
9-5	Number of DIF incidents completed by number of incidents eligible for a DIF and mode of interview for females ages 18-29 in the volunteer sample, 2014–2015	272
9-6	Sexual victimization screener question wording for ACASI.....	277
9-7	Percent of rapes and sexual assaults reported from SV screener items targeting.....	278
9-8	Prevalence rate of completed rape by type of crime classification, mode of interview and classification method for females ages 18-49 in the general population, 2014-2015.....	279

Contents (continued)

Tables		Page
9-9	Distribution of detailed incident form classifications among incidents classified by behaviorally specific questions as completed rape, by sample type and mode for females, 2014–2015	281
9-10	Distribution of behaviorally specific question classifications among incidents classified by detailed incident form as completed rape, by sample type and mode for females, 2014–2015	283
9-11	Prevalence rate of attempted penetration by mode of interview and classification method for females ages 18–49 in the general population, 2014–2015	293
9-12	Distribution of detailed incident form classifications among incidents classified by behaviorally specific questions as attempted penetration, by sample type, and mode for females, 2014–2015.....	294
9-13	Distribution of behaviorally specific question classifications among incidents classified by detailed incident form as attempted penetration, by sample type, and mode for females, 2014–2015.....	296
10-1	Primary prevalence rates from ACASI samples using intermediate and final weights, 2014–2015.....	308
10-2	Primary prevalence rates from CATI samples using intermediate and final weights, 2014–2015.....	309
10-3	Average percentages of tract-level characteristics of respondents (R) and nonrespondents (NR) for the ACASI sample, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015	311
10-4	Average percentages of tract-level characteristics of respondents (R) and nonrespondents (NR) for the ACASI sample, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015	312

Contents (continued)

Tables

	<u>Page</u>
10-5 Average percentages of county-level characteristics of respondents (R) and nonrespondents (NR) for the CATI sample, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015.....	314
10-6 Average percentages of county-level characteristics of respondents (R) and nonrespondents (NR) for the CATI sample, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015.....	315
10-7 Estimates of the proportion of outmovers and inmovers for ACASI and CATI samples, overall and by core-based statistical area, 2014–2015	319
10-8 Age distributions of the CATI cell phone outmovers and ACASI cell phone inmovers compared to cell phone nonmovers, 2014–2015	320
10-9 Prevalence rates of rape and sexual assault for CATI cell phone outmovers and ACASI cell phone inmovers compared to cell phone nonmovers, 2014–2015	321
10-10 Prevalence and incidence rates of sexual victimization by level of effort for ACASI respondents in the general population ages 18–49.....	324
10-11 Prevalence and incidence rates of sexual victimization by level of effort for CATI respondents in the general population ages 18–49.....	324
10-12 Coefficients for logistic regression analysis predicting prevalence of rape, sexual assault and other unwanted sexual contact, using level of effort and respondent characteristics, for females ages 18–49 in the general population, 2014–2015	326
11-1 Experimental conditions for the alcohol vignette	332
11-2 Percentage reporting that someone made Sue have sex when she was unable to consent because she was too drunk, high, or passed out, by consent, relationship type and drinking behavior for females ages 18–49 in the general population, 2014–2015	336

Contents (continued)

Tables

<u>Tables</u>	<u>Page</u>
11-3 Percentage reporting that a male used force or threats of force to make Sue have vaginal sex against her will by consent, type of relationship and drinking behavior for females ages 18-49 in the general population, 2014–2015.....	338
11-4 Percentage reporting that Tom was trying to get Sue drunk so he could sexually take advantage of her by consent, type of relationship and drinking behavior for females ages 18-49 in the general population, 2014–2015.....	340
11-5 Percentage reporting that alcohol made Sue unable to give consent by consent, relationship type and drinking behavior for females ages 18-49 in the general population, 2014–2015	341
11-6 Experimental conditions for the vignette involving verbal pressure	345
11-7 Percentage reporting that a male used force or threats of force to make Becky have vaginal sex against her will by use of gentle force, type of relationship and coercion for females ages 18-49 in the general population, 2014–2015	346
11-8 Percentage reporting that someone made Becky have sex against her will by threatening to cause problems for her, by use of gentle force, type of relationship and coercion for females ages 18-49 in the general population, 2014–2015.....	348
11-9 Type of event history calendar entry, by sample type, 2014–2015.....	351
11-10 Respondents' usage of the event history calendar during the interview, by sample type, 2014–2015	352
11-11 Regression coefficient for number of events entered on the event history calendar for equations predicting response time, amount of missing data, and external telescoping in the Incident dating questions for females ages 18 to 49 in the general population, 2014–2015.....	353
11-12 Victimization reports in the last 12 months, by number of event types entered on the event history calendar for females ages 18 to 49 in the general population, 2014–2015	355

Contents (continued)

Tables

	<u>Page</u>
11-13 Regression results predicting a past 12-month victimization based on the number of event types entered on the event history calendar for females ages 18 to 49 in the general population, 2014–2015.....	356
11-14 Parameter estimates and intraclass correlation coefficients from multi-level regression models predicting prevalence of rape, sexual assault and other unwanted sexual contact with respondent characteristics and interviewer experience, for females ages 18-49 in the general population, 2014–2015.....	368
12-1 Estimates of reliability for the sexual victimization screener items overall and by mode in the general population, 2014–2015.....	376
12-2 Type of sexual victimization screener incident reported in second interview, conditional on type of victimization reported in first interview, for the general population.....	378
12-3 Agreement rates and kappa statistic for detailed incident form question on unwanted vaginal penetration (D1a), 2014–2015	385
12-4 Agreement rates and kappa statistic for detailed incident form question on unwanted oral penetration (D1b), 2014–2015	386
12-5 Agreement rates and kappa statistic for detailed incident form question on unwanted anal penetration (D1c), 2014–2015	387
12-6 Agreement rates and kappa statistic for detailed incident form question on unwanted digital penetration (D1d), 2014–2015.....	388
12-7 Agreement rates and kappa statistic for detailed incident form question on unwanted kissing (D2a), 2014–2015	390
12-8 Agreement rates and kappa statistic for detailed incident form question on unwanted groping (D2b), 2014–2015	391
12-9 Agreement rates and kappa statistic for detailed incident form question on holding or pinning down (D4a), 2014–2015	392
12-10 Agreement rates and kappa statistic for detailed incident form question on whether passed out (G10), 2014–2015	394

Contents (continued)

Tables		Page
12-11	Agreement rates and kappa statistic for detailed incident form question on whether respondent can remember what happened (G11), 2014–2015	396
12-12	Agreement rates and kappa statistic for detailed incident form question on whether alcohol or drugs made respondent unable to give consent (G12a), 2014–2015	397
12-13	Agreement rates and kappa statistic for detailed incident form question whether the perpetrator immediately stopped after saying “no” (G17), 2014–2015	398
12-14	Estimates of reliability for detailed incident form classifications overall and by mode among the general population	401
12-15	Agreement rates and kappa statistic for respondents reporting victimization in the detailed incident form, for females in the general population sample, 2014–2015	403
12-16	Agreement rates and kappa statistic for detailed incident form classification of crime type at the first and second interviews, 2014–2015.....	407
13-1	Completed surveys by mode and sample type	411
13-2	Cost components by mode	412
13-3	RDD productivity comparison.....	415
13-4	RDD productivity by age group.....	416
14-1	Respondents’ reactions to the survey, by mode of interview for females ages 18–49 in the general population, 2014–2015	418
14-2	Respondents’ reactions to the survey, by mode of interview for females ages 18–29 in the volunteer sample, 2014–2015	419
14-3	Respondents’ reactions to the survey, by type of victimization for females ages 18–49 in the general population, 2014–2015	421

Contents (continued)

Tables

	<u>Page</u>
14-4 Respondents' reactions to the survey, by type of victimization for females ages 18-29 in the volunteer sample, 2014–2015	422
14-5 Respondents' reactions to the survey, by number of detailed incident forms completed for females ages 18-49 in the general population, 2014–2015	424
14-6 Respondents' reactions to the survey, by number of detailed incident forms completed for females ages 18-29 in the volunteer sample, 2014–2015	425
14-7 Respondents' ease of understanding the survey, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	426
14-8 Respondents' ease of understanding the survey, by type of victimization and mode of interview for females ages 18-29 in the volunteer sample, 2014–2015	427
14-9 Interviewer observations questionnaire items by mode.....	428
14-10 Interviewer ratings of distress, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015.....	430
14-11 Interviewer ratings of distress, by type of victimization and mode of interview for females ages 18-29 in the volunteer sample, 2014–2015.....	431
14-12 Interviewer ratings of cooperativeness and distress, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015	434
14-13 Interviewer ratings of cooperativeness and distress, by type of victimization and mode of interview for females ages 18-29 in the volunteer sample, 2014–2015.....	435
14-14 Interviewer ratings of the presence of others during the CATI interview, by type of victimization for females ages 18-49 in the general population, 2014–2015.....	438

Contents (continued)

Tables

	<u>Page</u>
14-15 Interviewer ratings of the presence of others during the CATI interview, by type of victimization for females ages 18 or older in the volunteer sample, 2014–2015.....	439
14-16 Interviewer ratings of the presence of others during ACASI, by type of victimization for females ages 18-49 in the general population, 2014–2015	440
14-17 Interviewer ratings of the presence of others during ACASI, by type of victimization for females ages 18-29 in the volunteer sample, 2014–2015	441
14-18 Interviewer ratings of what type of setting the ACASI interview was conducted in, by type of victimization for females ages 18-49 in the general population, 2014–2015.....	442
14-19 Interviewer ratings of what type of setting the ACASI interview was conducted in, by type of victimization for females ages 18-29 in the volunteer sample, 2014–2015	443
14-20 Interviewer ratings of respondent ACASI and event history calendar usage, by type of victimization for general population ages 18-49 (weighted).....	445
14-21 Interviewer ratings of respondent ACASI and event history calendar usage during ACASI, by type of victimization for volunteer sample (unweighted).....	446

Figures

E-1 Criteria used to define rape and sexual assault.....	xxxiv
2-1 Criteria used to define rape and sexual assault.....	24
4-1 Roster mailings and receipts.....	80
7-1 Comparison of mode effects in DIF prevalence estimates between general population and volunteer samples by victimization type for females age 18-29.....	156

Contents (continued)

<u>Figures</u>	<u>Page</u>
9-1 Recency distribution of incident dates reported in the ACASI interview mode by type of victimization for females age 18-49 in the general population.....	261
9-2 Recency distribution of incident dates reported in the CATI interview mode by type of victimization for females age 18-49 in the general population.....	262
9-3 Recency distribution of incident dates reported in the ACASI interview mode by type of victimization for females age 18-29 in the volunteer sample	263
9-4 Recency distribution of incident dates reported in the CATI interview mode by type of victimization for females age 18-29 in the volunteer sample	264
11-1 Percent of adequate answers by victimization screener incident dating items for females 18+ in the general population (weighted).....	362
11-2 Percent of adequate answers by victimization screener incident dating items for females age 18-29 in the volunteer sample (unweighted).....	365
12-1 Reasons for changing answers to 12-month victimization screener items (ACASI)*	379
12-2 Reasons for changing answers to lifetime victimization screener items (CATI)*	380
14-1 Range of respondents in the GP sample and prior research studies that reported negative and positive reactions.....	420

Contents (continued)

<u>Exhibits</u>		<u>Page</u>
<u>Exhibit 11-1.</u>	<u>Event history calendar</u>	350
11-1	Event history calendar	350
15-1	Recommendations for collecting data on rape and sexual assault within the NCVS data collection program.....	464
15-2	Example of revised sequence to measure behaviors and tactics	482

Disclaimer and Acknowledgments

Disclaimer

Findings, opinions, or points of view expressed in this report are those of the author(s) and do not necessarily reflect official findings, positions or policies of the Bureau of Justice Statistics or the U.S. Department of Justice. The results of the pilot test are not intended to provide official estimates of the prevalence or incidence of rape or sexual assault in the five metropolitan areas selected for the study.

Acknowledgments

The authors wish to express gratitude to the many individuals who contributed to the successful execution of this research. First and foremost, we thank the women who agreed to be interviewed and provide sensitive information about their personal experiences. We thank all of the victim service organizations that agreed to help facilitate making contact with women who were receiving services at the time of the survey.

Shannan Catalano, the BJS project manager, and Allen J. Beck, BJS's senior statistical advisor, provided critical insights into key methodological and substantive issues. They collaborated with the Westat team on all phases, including sample design, recruitment, questionnaire design, data collection and analysis.

We thank the members of the CNSTAT Panel on *Estimating the Incidence of Rape and Sexual Assault* for providing a forum to discuss the issues related to measuring rape and sexual assault. We thank André B. Rosay for providing the estimates from the *National Intimate Partner and Sexual Violence Survey* that are included in Chapter 7.

Several organizations and individuals provided assistance when reviewing and classifying incidents during the analysis. Jennifer Long and Charlie Whitman-Barr of AEQUITAS provided several rounds of feedback on different aspects on the definitions of rape and sexual assault for the study. Virginia Baran, Suzanne Pugliese, and Marnie Shield of the Office of Victims of Crime also provided feedback on the legal underpinnings of different definitions.

We would like to thank the large team of individuals at Westat who helped carry out this research. Chris de los Santos, Patrick Mayer, Drew Kistler, Rebecca Kirkner, and Stacia Noble led the development of the survey programs. A number of individuals assisted in the development of the questionnaires, including Wendy Hicks, Michele Harmon, Eileen Ahlen, Teresa Koenig, Jerelyn Bouic, Gina Shkodriani, Cecilia Avison, Jennifer Anderson, and Cristina Golab. Kerry Levin and Sharon Zack provided guidance working with the IRB. For the in-person data collection, Brad Edwards provided guidance on the design of the procedures. Carol Cober and Sherry Sanborne developed the training materials and Cindy Good led the recruitment of universities to interview college students. Amanda Hall and her team of field supervisors, field managers, and field interviewers carried out the collection. The telephone data collection was led by Sarah Dipko and was carried out by her team of supervisors and telephone interviewers. The analysts included Andrew Caporaso, Rose MacAloon, Xiaoshu Zhu, Esteban Gonzalo, Mariel Leonard, Trey Arthur, Bryan Williams, Julia White, Terrell Hayes, and Kay Ricci. Alfred Bishop led the data programmers and Reina Sprankle led the production support. Bruce Herzig, Linda Allers, and Debra Reames provided editorial and word processing support.

Executive Summary

Over the past two decades, data collections using different methodologies and definitions have generated competing estimates on the level of sexual victimization in the United States. Estimates based on the National Crime Victimization Survey (NCVS), conducted by the Bureau of Justice Statistics (BJS), have been lower than estimates obtained from surveys administered by other Federal agencies and by private groups (Black et al., 2011; Koss & Gidycz, 1985; Tjaden & Thoennes, 2000; Kilpatrick, 2007; Fisher et al. 2000). For example, 2010 estimates of rape from the National Intimate Partner and Sexual Violence Survey (NISVS) (Black et al., 2011) were more than 10 times higher than estimates of rape and sexual assault from the 2010 NCVS (Truman, 2011). These differences have led to confusion on the level of rape and sexual assault in the nation (e.g., Gilbert, 1997; Lynch, 1996; Rand & Rennison, 2005; Bialik, 2013).

BJS initiated the Rape and Sexual Assault Pilot Test (RSA Pilot Test) to develop recommendations on the best methods to collect data on rape and sexual assault within the NCVS program. The project had three objectives:

1. to develop and pilot test a design based on best practices to collect self-report data on rape and sexual assault using an in-person, audio computer-assisted self interview (ACASI) questionnaire
2. to develop and pilot test a comparison design using random digit dialing (RDD) and a computer-assisted telephone interview (CATI) survey
3. to conduct detailed analytical comparisons of the two designs against each other and the existing NCVS program.

The ACASI design was intended to implement procedures that maximized the overall survey response and coverage rates using a self-administered interviewing mode. This mode has been found to be best for collection of sensitive data. The CATI design provided a comparison group that used a sample frame and mode that has been used on prior surveys collecting rape and sexual assault. When comparing the two surveys with each other and the NCVS, the project evaluated the accuracy, utility, and cost of each of the methods.

One goal of the project was to improve the data collection methodology and measurement within the NCVS program. The RSA Pilot Test developed recommendations for the NCVS program by testing several different designs, which implemented features that were thought to improve data quality relative to the NCVS. As part of this goal, the project was to determine

whether it was possible to accommodate improvements in measurement within the NCVS or whether a separate survey was needed.

A second goal was to contribute to knowledge on the best methods to collect data on rape and sexual assault. There have been several reviews of methods used to conduct victimization surveys collecting rape and sexual assault data, including why estimates differ between the NCVS and other surveys (Lynch, 1986; Cook et al., 2011). Five issues were investigated:

1. How and why do estimates from a survey using behaviorally specific questions differ from the NCVS?
2. Is there a difference in estimates of rape and sexual assault between self-administered and interviewer-administered modes?
3. Are there significant effects of non-response bias for a survey on rape and sexual assault for either the in-person ACASI and RDD CATI surveys?
4. What are the advantages and disadvantages of a one-stage vs. two-stage design?
5. What is the data quality (validity and reliability) of estimates from a questionnaire using behaviorally specific questions (BSQs)?

This executive summary reviews the methodology, key findings, recommendations, and limitations of the study. A full summary of the conclusions and recommendations can be found in Chapter 15 of the report.

Methodology

Two surveys were administered. One survey was administered with an audio computer-assisted self-interview (ACASI) via in-person visits. The second survey was administered with a computer-assisted telephone interview (CATI). The two surveys were nearly identical in content and structure, with two exceptions. The CATI interview restricted most questions to a “yes/no” response format to preserve confidentiality within the household. A second difference was how victimization over the lifetime and the last 12 months were administered. In CATI, questions about lifetime victimization were administered first, while in ACASI, last 12 month victimizations were administered before lifetime questions (see Chapter 2).

A two-stage questionnaire design was used to measure four types of sexual violence, comprising rape, sexual assault, coercive contact, and unwanted contact (figure E-1). At the first stage, a series of 14 BSQs asked about different combinations of behaviors and tactics. Once specific incidents were enumerated from the BSQs, a detailed incident form (DIF) was used to collect information on each incident. Just as with the NCVS, the final classification of the incident into rape, sexual assault, coercive contact, and unwanted contact was based on responses to the DIF, as well as information from narratives provided by the respondents.

A rape was defined as one of four different types of completed, attempted, or threatened penetration (vaginal, anal, oral, or digital) that occurred without the respondent's consent because of physical force or because the respondent was unable to consent. Sexual assault was defined as completed, attempted, or threatened non-penetrative sexual contact (e.g., kissing, groping, touching) or non-contact (e.g., exposure, exploitation by photos or video) by physical force or inability to consent. Coercive acts included sexual contacts where the offender threatened non-physical punishment (e.g., threats of job loss) or offered rewards (e.g., financial support, better grades). Unwanted sexual contact was defined as any other type of sexual contact the respondent said she did not want to do.

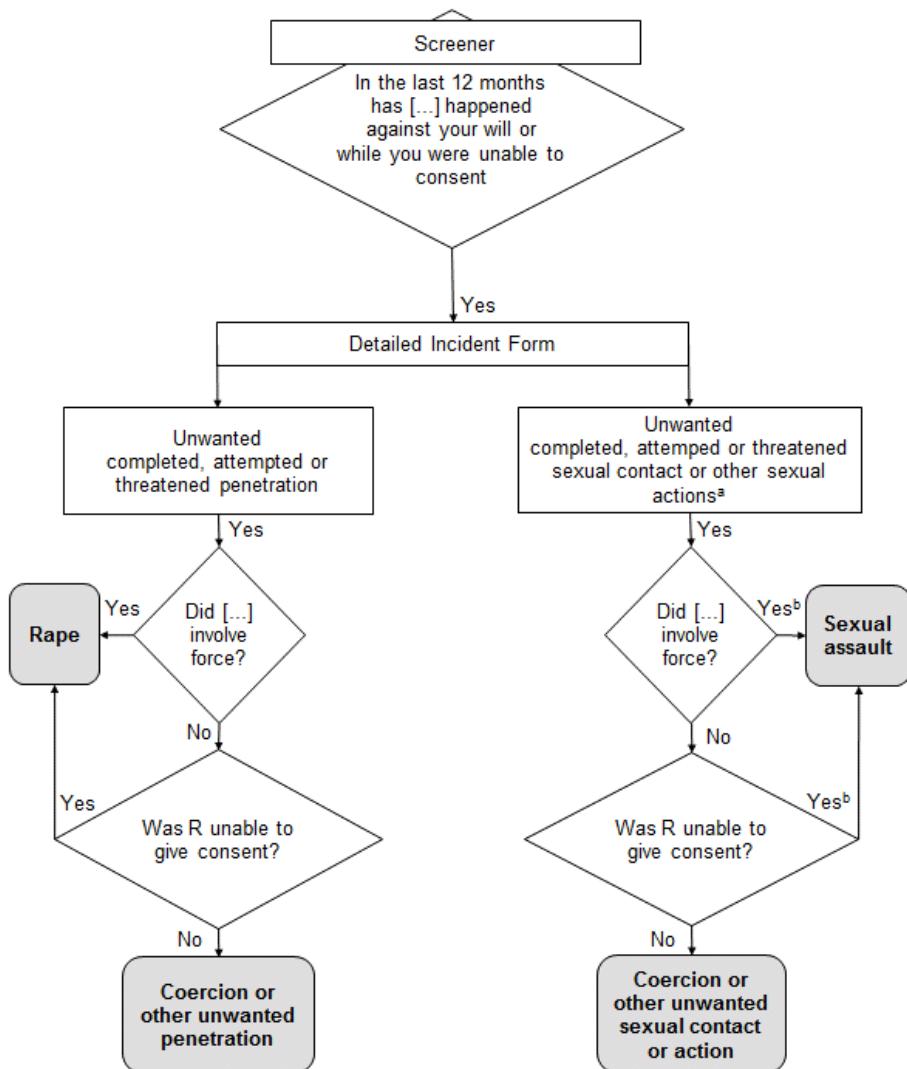
Three types of samples were used to interview English- and Spanish-speaking women in five large core-based statistical areas (CBSAs): Dallas, Los Angeles, Miami, New York City, and Phoenix. One type of sample was for the general population. Two sample frames were used for this: (1) an address-based sample was drawn to interview respondents using ACASI and (2) a random digit dialing (RDD) frame of landline and cell phone numbers was used to sample and interview respondents using CATI. For the general population samples, the primary analysis was among women age 18 to 49.

The second type of sample was a group of women age 18 to 29 who volunteered to participate in the study via a Craigslist advertisement. The third type of sample was composed of volunteers who were clients at rape crisis centers in each MSA. Volunteers from the Craigslist and rape crisis centers were randomly assigned to either the ACASI or CATI mode of interviewing (for more information on types of samples, see Chapter 2).

Approximately 11,000 interviews were completed across the three sample types and the two modes of interviewing. For the general population sample, the response rate for the ACASI was 40

percent and for the CATI 18 percent (AAPOR RR2).¹ Evaluation of non-response bias in the two modes used Census data, as well as comparison of responses by the level of effort to complete the interviews. Data quality was evaluated by various methods including information from the detailed incident form, approximately 1,000 re-interviews, vignettes administered at the end of the interview, debriefing information provided by respondents, and interviewer observations.

Figure E-1. Criteria used to define rape and sexual assault



^aIncludes threatened sexual contact, exposure or video.

^bIf the offender stopped immediately, the incident was not classified as sexual assault.

¹ https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions2015_8theditionwithchanges_April2015_logo.pdf, page 52

Key Findings – Data Quality

- 1. No significant non-response bias was found for the RSA Pilot Test estimates. With respect to coverage, the random digit dial sample frame resulted in undercoverage of rape victims.**

One question related to data quality is whether the estimates from one or both of the surveys had significant non-response bias. Non-response bias is one source of error that has been discussed when comparing the NCVS, which has a relatively high response rate, to many of the RDD surveys, which generally have lower response rates. To examine this question, the RSA Pilot Test compared the effects of non-response on the ACASI and CATI estimates. The ACASI was based on in-person contact and achieved a response rate of approximately 40 percent. The CATI survey was based on a random digit dial sample frame and had a response rate of 18 percent (Chapter 3). Analyses did not find evidence of significant non-response bias in either survey. With respect to coverage error, using Census characteristics for the areas where the respondents were located showed no indications of significant under- or overrepresentation of demographic or economic groups for the surveys conducted with either mode of interview (Chapter 10). A level of effort analysis did not find a correlation between the number of contacts needed to complete the survey and reporting victimization. These conclusions are consistent with several recent campus climate surveys (Cantor et al., 2016; Krebs et al., 2016), as well as early analysis of the NISVS (Peytchev et al., 2009). These surveys did not find evidence of significant non-response bias for outcomes related to rape or sexual assault among surveys with response rates between 19 percent and 40 percent.

As with most evaluations of non-response, the above analysis cannot definitively conclude that estimates from either the ACASI or CATI data collections are free of bias. The evidence from the level-of-effort analyses, cited above, are based on the assumption that the late responders to the RSA Pilot Test surveys are representative of the non-responders. In particular, the analysis assumes that the 60 percent of the sample that did not respond to the ACASI and the 82 percent that did not respond to the CATI survey have similar rates of rape and sexual assault as the late responders to the survey. There was no way to test this assumption with the data from this study. Several studies of other types of outcomes have shown that this assumption does not always hold (Lin & Shaeffer, 1995). Future research, perhaps with surveys that have a higher response rate (e.g., the NCVS) or have data on non-responders, would shed more light on the role non-response may play on surveys of rape and sexual assault.

The CATI survey was based on an RDD sample frame, which did not allow precise targeting of sampled cases within the sampled CBSAs. The frame is based on area codes and there are a significant number of individuals with mobile phones who have area codes from outside the CBSA. For example, approximately 10 percent of the ACASI respondents who owned a mobile phone did not have an area code that was included on the RDD frame. Analysis of this group of respondents found them to have significantly higher rates of rape than the general population.

- 2. For most items, item missing data rates were low. There were higher rates of missing detailed incident forms when they were requested for multiple incidents. This was higher for the CATI than the ACASI. Interviewer variance related to the CATI interview was within the range of other telephone surveys.**

With a few exceptions, item-missing data were generally low (below 5 percent) for both modes of interviewing. The highest missing data rate was associated with the narrative collected at the end of each detailed incident form, with about 30 percent of incidents missing a narrative. Missing data were also a problem for filling out multiple detailed incident forms. Almost all respondents completed a detailed incident form for one incident. However, the willingness of respondents with multiple victimizations to complete all of the assigned detailed incident forms decreased as the number of incidents increased (Chapter 9). The rate of missing incident forms was significantly higher for the CATI interviews. This last result was attributed to the additional time it took to administer the detailed interview form for this interview mode, which required yes/no responses.

For the CATI interviews, there were no large interviewer effects associated with administering the survey (Chapter 11). The intra-class correlation associated with the interviewers was not statistically significant and was of a magnitude similar to other CATI studies.

- 3. With respect to the estimates of victimization, there were very few differences between ACASI and CATI interviewing modes.**

As described in several publications (e.g., Kruttschnitt, 2014; Tourangeau & Yan, 2007), it was expected that the increased anonymity of a self-administered survey mode, such as ACASI, would produce higher estimates than an interviewer-administered survey mode, like CATI. The ACASI mode in this study did produce nominally higher estimates of rape and sexual assault within the last 12 months than the CATI mode. The incidence estimate for rape for the ACASI design was

51.0 per 1,000 women age 18-49 compared to 43.1 for the interviews conducted using the CATI (see Chapter 7). The estimates of prevalence, expressed as the percentage of women age 18 to 49 who had experienced rape or sexual assault in the prior 12-month period, were 5.9 for ACASI and 5.3 for CATI. However, neither of these differences were statistically significant. Furthermore, there is evidence that the nominal difference is related to differential coverage of the address-based sampling (ABS) and RDD frames (see Finding 1 above; Chapter 10), as well as higher external telescoping in the ACASI (Chapter 9). There was some indication that mode has a greater effect among individuals who experience more incidents. For the sample of volunteers who had the highest risk of victimization, there was a significant effect of mode, in the expected direction (i.e., higher rates for ACASI).

With a few isolated exceptions, the nature and characteristics of the reported incidents also did not differ by mode (see Chapter 8). For example, the percentage of incidents in which the offender is known to the victim is the same across modes. The extent of injuries that occur, when the incidents happen, where they occur, and whether the police are notified about the incident were also the same across modes.

- 4. *The overall prevalence rate for rape just based on the BSQs was similar to the rate based on the DIF classification. However, there was a significant difference in which particular incidents were classified as a rape using the two methods. Both the BSQ and the DIF are subject to measurement error. Collecting a narrative of what happened provided a useful supplement to evaluate the final classification.***

One reason to use a two-stage method is to provide detailed information about the characteristics, circumstances, and consequences of the events. A second reason is to collect the information needed to classify an incident into particular types of crimes and to count unique incidents so that incidence estimates can be generated. Prior specialized surveys measuring RSA have used the victimization screener primarily to classify events and have concentrated on prevalence estimates, which do not require identification of individual incidents.

For completed, attempted, and threatened rape, the RSA Pilot Test compared classification of the incident using only the screener questions to the classification using the DIF. Comparing the two classifications and resolving differences using the narratives provided a way to assess the accuracy of each classification methodology (Chapter 9).

The prevalence rates using the BSQs and those using the DIF were very similar to one another. For example, the ACASI prevalence of completed rape from the BSQ was 2.8 percent compared to 2.4 percent based on the DIF. The rates were almost identical for the CATI estimates. Similar results were found for the measurement of prevalence of attempted rape. However, there were quite a few discrepancies between how particular incidents were classified. Of the ACASI incidents classified by the detailed incident form as a completed rape, 30 percent came from a BSQ that did not target this type of crime. Conversely, of the ACASI incidents classified as a completed rape by the BSQ, 50 percent were not classified as such by the detailed incident form.

Using the narratives to resolve discrepancies, measurement error was found for both the BSQs and the DIF. With respect to the BSQs, some RSA Pilot Test respondents answered affirmatively to a BSQ when their experience includes some, but not all, elements stated in the question. Other respondents reported a rape in response to questions that were not targeting this type of crime. This may occur because the respondent remembers the incident at that particular BSQ or decides to report an incident after thinking about it at a later point in the survey. Measurement error was also found on the DIF. One source of error was centered on the questions determining unwanted behaviors, which attempted to make distinctions among threats, attempts, and completed acts. This level of detail seem to have confused some respondents when trying to describe what happened because respondents tended to have a wider definition of attempt or a threat than intended. Another source of error were related to the questions on physical force. The list of alternative types of force presented to the respondent was missing tactics related to some types of groping incidents.

- 5. Data reliability of the BSQs and of portions of the DIF was lower than expected for behavioral data. Incidents involving attempts and threats were the most unreliable. For the BSQs, respondents reported re-interpreting the questions at the second interview, remembering incidents that were forgotten between interviews, and not reporting incidents to avoid the follow-up questions. For the DIF, the questions related to unwanted behavior were one source of inconsistency. The second source was the omission of a response category in the item on physical force for groping incidents.**

Data reliability was examined by conducting approximately 1,000 re-interviews, two to three weeks later, among those who reported an unwanted sexual contact at the first interview. The screener items, when grouped into logical categories, had Kappas of 60 to 70 percent, depending on which mode is examined; these are considered to be “substantial” using the Koch and Landis (1977) standard. Nonetheless, this is lower than one might expect given the fairly specific behaviors that are

being referenced on the screener. For example, approximately 30 percent of respondents who reported a completed rape on a screener item at the first interview changed answers to something else at the second interview. The change was distributed across the remaining types of victimizations, including not reporting any 12-month incident at all. When respondents were asked why they changed responses, the most prominent reason was re-interpretation of the questions at the second interview. The next most common reason was something being remembered differently at the second interview. Some respondents said they did not report an incident to avoid being asked the follow-up questions.

The analysis of the DIF concentrated on those incidents that could be identified as being the same based on the narrative. This provided a direct measure of whether respondents were answering the DIF items the same way for the same set of target behaviors across the two interviews. Based on this subset of incidents, the questions on unwanted behaviors had lower reliabilities than expected. Some of this was related to confusion about how to define an attempted or threatened act. There also were higher than expected inconsistencies related to respondents reporting a completed act (e.g., penetration, sexual touching). Most of the latter were respondents reporting a completed act at the first interview and no unwanted behavior at the second interview. For example, of those that reported a completed vaginal penetration, 30 percent reported that no unwanted vaginal penetration occurred. The other items used to classify the incident, such as “using force” and “inability to consent,” had relatively high agreement rates and reliabilities.

When examining the consistency of the overall crime classification algorithm, which combined items to determine the type of crime, completed rape had the highest consistency of the major categories, with approximately 70 percent of the incidents being classified the same way at both interviews. Consistent with the discussion above, many of the changes in classification between interviews were related to respondents reporting completed penetration at one interview and no type of penetration at the second interview. Review of the narratives suggests that this type of change was due to problems with the structure of the unwanted behavior questions, as the narratives indicated penetration had occurred in most cases. Sexual assault had the lowest consistency. For completed sexual assaults, the change was related to the omission of a category for groping and grabbing in the force question. The attempted and threatened acts of sexual assault were subject to a similar problem; however, there was also some indication that change between interviews was related to ambiguity with respect to what constituted a forced attempt or threat.

- 6. Based on a review of the narratives, the incidents classified as rape and sexual assault generally conformed to the intended definitions. There was little indication that the prevalence and incidence rates were subject to a high number of false positives. There is no clear method for determining false negatives in this study.**

The use of BSQs and multiple screening items, as well as increasing the privacy of the interview, are intended to reduce false negatives (underreporting). One possible problem with this methodology is that this may come at the cost of increasing false positives (overreporting). A false positive is an incident reported on the survey that is outside of the scope of the survey in terms of the reference period or type of behavior (table E-1). By asking multiple questions about related behavior, respondents may feel pressure to answer in the affirmative. For example, the respondent may report an incident that occurred outside the reference period or may include an incident even though they are unsure whether it meets the criteria related to the question (e.g., does not involve clear-cut use of force or inability to consent).

One purpose of a two-stage design is to evaluate whether specific incidents meet the operational criteria used to define a rape or sexual assault. For each incident reported on the RSA Pilot Test, a series of questions were asked about the use of physical force and the respondent's ability to provide consent. These are the two elements that define an incident as a crime. With respect to physical force, about 80 percent of incidents that were classified as a completed rape included some type of physical force, such as being held or pinned, the use of a weapon or threat with a weapon, a physical attack or threat of physical attack without a weapon,² a physical attack or threat to attack someone else, or being blocked or otherwise prevented from leaving (e.g., locked in, handcuffed).

Table E-1. Definition of false negative and false positive errors

Survey measure	Truth	
	Not victimized	Victimized
Not victimized	<i>True negative</i>	<i>False negative</i>
Victimized	<i>False positive</i>	<i>True positive</i>

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

² This includes a “yes” response to a direct question of whether physical force was used, as well as the respondent saying she was hit, punched, bitten, choked, slapped, kicked, had her mouth covered, grabbed, pushed, pulled, or groped.

In the RSA Pilot Test, about 20 percent of rape incidents involved no physical force, but the victim was classified as being unable to consent.³ Incidents involving inability to consent are sometimes difficult to prove in a court proceeding and pose unique problems with respect to measuring on a survey. Most who met the criteria for inability to consent also reported signs of being intoxicated to the point they had trouble making decisions and/or provided an indication that another individual would recognize the victim was not in condition to consent.⁴ Slightly more than half of the incidents classified as “rape due to inability to consent” involved victims who were unconscious for at least part of the incident. In slightly less than half of the incidents, the victim was conscious but said she was unable to consent because of alcohol or drugs. Approximately 90 percent of these victims also reported one or more additional signs of inability to consent, such as not being able to communicate with others, not being able to walk by herself, being less able to physically resist, and having the perpetrator continue to give her alcohol or drugs after she was clearly drunk or high (Chapter 8).

The RSA Pilot Test also included questions about how victims reacted during the incident. These items provide a picture of how the victim expressed non-consent during the incident. For incidents involving kissing, groping, or other type of sexual touching, if the respondent reported the perpetrator immediately stopped after she said “no,” the incident was not counted as a sexual assault, regardless of the tactic or ability to consent. These data also reveal that the vast majority of the victims of rape and sexual assault did express non-consent during the encounter. Of the incidents classified as a rape, 84 percent of victims physically resisted or tried to physically resist,⁵ 92 percent said “no” or “stop,” 68 percent pleaded or argued with the perpetrator, and 64 percent tried to escape or get away.

A final piece of evidence related to false positives is a review of narratives provided by the victims about what happened. About 70 percent of the respondents across both modes provided narratives. These descriptions were reviewed and compared to the RSA Pilot Test survey responses to assess the accuracy of the information provided on the structured instrument. Generally speaking, these descriptions were consistent with the information used in the classification of the incident as a completed rape or sexual assault (see Chapter 7). While some false positives were found, there were relatively few. The classification into attempted or threatened rape was more problematic.

³ There were a number of incidents where both force and inability to consent occurred. The study counted these incidents as involving force.

⁴ This is a key element in prosecuting alcohol/drug-related rape cases.

⁵ These refer to the results for the ACASI survey. The results for the CATI are similar.

Respondents tended to have a broader definition of what constituted an attempt or threat than the legal criteria. The classification scheme used in the analysis made adjustments for these incidents when the narrative provided enough detail (Chapter 7). Chapter 15 makes recommendations on how to improve the survey items that ask about attempts and threats.

7. *The RSA Pilot Test procedures successfully maintained the privacy of the interview as well as mitigating harm related to emotional reactions associated with reporting rape and sexual assault.*

There were several human subject concerns addressed on the project (Chapter 14). The first was to implement procedures to mitigate harm to respondents who may have adverse reactions to the survey. Prior research has found that respondents do not report substantial harm from taking sensitive surveys similar to the RSA Pilot Test. Even those who do report negative feelings do not typically regret taking the survey because they also report experiencing positive emotions such as empowerment or contributing to solving the problem (DePrince & Chu, 2008; McClinton et al., 2015; Newman et al., 2006). One unique feature of the RSA Pilot Test is administration of a relatively lengthy detailed incident form, which could be perceived as more intrusive than prior surveys. Findings from the RSA Pilot Test are in line with prior research, with rates for positive and negative reactions falling within ranges reported elsewhere (Black et al., 2006; Valpied et al., 2014; Wager, 2012; Walker et al., 1997). Respondents in the RSA Pilot Test reported higher agreement with items indicating more positive reactions than negative reactions, and most respondents indicated that they did not regret taking the survey. This general pattern held true across survey modes and levels of victimization.

Similarly, interviewers noted that very few respondents exhibited signs of distress at any point during the interview – around 2.5 percent regardless of sample type. The distress that was noted was found to be low and did not lead to significant breaks in the interview. None of the interviews resulted in a high level of distress, which would have required stopping the interview.

A second human subject concern was being able to conduct the interview in an environment that preserved the confidentiality of the interview. Both the ACASI and CATI interviews were administered to maintain this privacy. In both cases, the topic of the survey was not revealed until administering the informed consent to the selected respondent. For the in-person visit, the informed consent was incorporated on the ACASI. For the telephone interview, most questions were structured so respondents only had to answer “yes” or “no” to make it difficult for anyone else in the household to understand what the interview was about. One concern for the ACASI was

avoiding situations where someone else in the household was able to hear or see the interview. Especially in small housing units (e.g., small apartments), it may not be possible to fully prevent individuals from walking into the interviewing space or overhearing from another room. Isolating respondents in a private setting may be difficult to negotiate in a household setting. For the CATI, the interviewer had less control over who might be listening to their interactions.

The survey was successful in implementing the privacy and confidentiality conditions for the interview. A very small percentage of the CATI interviews were done where the interviewer suspected that someone was listening in on any part of the interview. For ACASI, in about one-third of the interviews, someone entered the room at some point during the interview. About 60 percent of the time when someone entered the room, it was a child. The interviewers were successful in maintaining the privacy of the ACASI questions. A very small percentage of the interviews occurred where someone was looking over the shoulder of the respondent at any time during the interview (1.8%). The interviewers were instructed to stop the interview when this occurred. Home office staff followed up in all instances to verify the interviewers took the appropriate action when this occurred.

Key Findings – Comparison to the NCVS

- 8. RSA Pilot Test estimates are substantially higher than NCVS estimates. One primary source of the difference is the expanded set of questions included on the RSA Pilot Test measuring a wider scope of incidents. Other possible sources of differences include the greater number of screening questions on the RSA Pilot Test, the criminal justice focus of the NCVS, and the privacy afforded by the RSA Pilot Test procedures.**

The two RSA Pilot Test surveys produced significantly higher incidence estimates than those from the NCVS. Among women age 18 to 49 in the general population in the five CBSAs, the NCVS incidence rate for rape and sexual assault was 1.5 per 1,000 women (see Chapter 7). The RSA Pilot Test rates are approximately 50 times higher than the NCVS for this age group. When focusing on just rape, the RSA Pilot Test estimates are 60 to 70 times higher. (See table E-2.)

Table E-2. Incidence rate per 1,000 of rape and sexual assault comparing NCVS to RSA Pilot Test for females ages 18-49 in the 5 CBSAs

	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Rape and sexual assault ^e	1.5	81.0 †	84.9 †	77.0 †
Rape ^f	0.7	47.1 †	51.0 †	43.1 †
Sexual assault ^g	0.8	33.9 †	33.9 †	33.9 †

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^aBased on NCVS for the years 2011-2014 among women ages 18-49 living in the 5 CBSAs of the RSA Pilot Test.

^bBased on females ages 18-49 in the general population sample of the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^fIncludes penetrative sexual contact using force or while unable to consent.

^gIncludes non-penetrative sexual contact using force or while unable to consent. Note the NCVS estimate also includes unwanted sexual contact without force.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2011-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

A large portion of the difference is attributed to the scope of incidents that are measured through the two collections. The NCVS was last redesigned in 1992 when definitions and social response to rape and sexual assault were rapidly evolving. The NCVS survey items explicitly use the term “rape” without providing any clarification that rape can include other types of penetrative acts beyond penile-vaginal penetration (digital, oral, anal). In addition, the NCVS focuses on incidents involving force, using terms like “attack” and “force,” which may discourage respondents from reporting acts involving unwanted grabbing or touching or attempted acts. The NCVS also does not ask directly about incidents involving an inability to consent (e.g., due to alcohol or drugs). The BSQs used in the RSA Pilot Test included vaginal, oral, anal, and digital penetration using force, any form of penetration while unable to consent, attempted penetration using force or while unable to consent, and acts involving unwanted touching or grabbing.

To control for the differences in scope between the RSA Pilot Test and the NCVS, estimates were generated for the RSA Pilot Test using BSQs that more closely align with the NCVS, which focuses on penile-vaginal penetration and forced completed acts. This reduced the difference

between the RSA Pilot Test and the NCVS rates by more than two-thirds (27 victimizations per 1,000 for the RSA Pilot Test vs. 1.5 per 1,000 for the NCVS).

A second difference between the collections is the framing of the survey context for the respondents. As a crime survey, the NCVS cannot mask the context in which respondents are asked to report their experiences. Whether BSQs are introduced or current screening items are enhanced, the scope of the incidents that are measured in the NCVS and in surveys related to health and safety may differ just for this reason. The NCVS has “crime” in its title and uses the word “crime” in several other key points throughout the screener.⁶ The questions on rape and sexual assault are also preceded by other types of predatory acts that are commonly considered criminal (e.g., robbery, burglary, motor vehicle theft). In contrast, the RSA Pilot Test was introduced as a survey about health and safety, and the questions do not refer to any acts as being criminal. On the RSA Pilot Test, 40 percent of the victims of rape and sexual assault thought the incident was a crime at the time it occurred. While no comparable data on the NCVS data exist, one could hypothesize that NCVS respondents are cued to recall incidents they believe are crimes.⁷ To account for the differences in the ways the two surveys are framed, estimates were generated from incidents on the RSA Pilot Test that the respondent considered to be a crime. This adjustment, in addition to the prior adjustment for type of crime, reduced the difference between the NCVS and the RSA Pilot Test by another 75 percent (7.3 per 1,000 for the RSA Pilot Test vs. 1.5 per 1,000 for the NCVS).

There are several other possible reasons why the RSA Pilot Test and NCVS estimates differ that are more difficult to quantify. One is the privacy of the interview. Everyone in the household age 12 and over is interviewed on the NCVS and thus everyone knows what questions are asked on the survey. While NCVS interviewers are trained to try to conduct interviews in private, both in-person and telephone interviews may be conducted within earshot of other household members. The RSA Pilot Test survey, in contrast, interviewed only one person per household and did not reveal the topic of the survey to any other member of the household. For the in-person visits, the interview was conducted using ACASI, so no other household member could hear any exchanges related to the topic or the survey items. Telephone survey respondents were encouraged to stop the interview if they thought someone in the household might be listening.

⁶ For example, at the beginning of the NCVS victimization screening questions, the respondent is instructed, “I’m going to read some examples that will give you an idea of the kinds of crimes this study covers.”

⁷ See Chapter 1 for a discussion of the issues when relying on incidents that respondents believe are crimes to measure rape and sexual assault.

Another difference is a greater number of screening questions specifically asking about sexual victimization on the RSA Pilot Test compared to the NCVS (14 vs. 2). Survey methodologists have found that asking more questions that mention the targeted or related behaviors will produce higher rates of reporting. This was the basis of the redesign of the NCVS in 1992, which increased the measured victimization rates by increasing the density of the relevant cues (Bidnerman et al., 1986).

Although the NCVS includes a number of screeners pertaining to general criminal victimization, exposure to multiple questions on different types of sexual violence specifically could have triggered memory of incidents that would not have otherwise been reported. Indeed, some respondents reported incidents after passing a targeted item earlier in the survey questionnaire. As noted above, the RSA Pilot Test included 14 different screening items specific to sexual victimization. Five of these questions were selected to enumerate specific behaviors and tactics that constitute rape. These questions elicited incidents that accounted for about 80 percent of the rapes reported on the survey (table E-3). The remaining 20 percent were reported at BSQs that asked about other types of behaviors and tactics. Approximately 75 percent of sexual assaults were reported at three different items targeting non-penetrative sexual contact and non-contact and 25 percent were reported at screener items that mentioned a different behavior. As noted above, the NCVS includes two questions that directly ask about rape and sexual assault.

Table E-3. Percent of rapes and sexual assaults reported from BSQ items targeting different behaviors and tactics

BSQ items targeting	# of items	Percent of:	
		Rapes	Sexual assaults
Penetration against victim's will by force or unable to consent	5	54%	5%
Other penetration against victim's will	2	5%	1%
Attempted penetration against victim's will	1	29%	16%
Kissing, sexual touching against victim's will	2	7%	53%
Attempted touching against victim's will	1	3%	22%
Exposure, photo or videos against victim's will	3	2%	3%
Total	14	100%	100%

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Some caution should be taken when interpreting the above differences. The NCVS and the RSA Pilot Test were completed with different designs and by two different organizations. The RSA Pilot Test was conducted as a one-time survey, conducted by a private contractor. The NCVS is conducted by the Census Bureau, an agency of the Federal government, as part of a rotating panel

design. The panel design may contribute to lower estimates on the NCVS because responses are not subject to telescoping and respondent conditioning. For example, the NCVS downweights the first interview by factors as low as .5 to account for bounding and panel conditioning. The NCVS also has a significantly higher response rate than the RSA Pilot Test. In the particular cities that were included in the RSA Pilot Test, the average response rate on the NCVS was approximately 28 percentage points higher than the RSA ACASI survey and 50 percentage points higher than the RSA CATI survey. If non-victims are more likely to participate on the NCVS relative to the RSA Pilot Test, then some of the difference between the estimates may be related to non-response bias.

- 9. *The types of victims the RSA Pilot Test and the NCVS are similar. While the types of incidents reported on the two surveys have some commonalities, there are more differences. One of the major differences is the NCVS exhibits a higher percentage of incidents for which the police were contacted. Other differences are related to the relationship to the perpetrator and the consequences of the incident.***

The correlates of victimization risk were very similar for the two studies. For the RSA Pilot Test, the strongest correlates associated with both rape and sexual assault were age, marital status, and race/Hispanic origin. Females ages 18 to 24 had victimization rates for rape that were significantly higher than those just a few years older (25-29); this trend continued into the older age groups. Those who are married had significantly lower rates than those not married. With respect to race/Hispanic origin, non-Hispanic white women have the highest rates. Household income was also significantly related to rates of rape. Those in the lowest income group had the highest victimization rates. Analysis of the NCVS, both in bivariate (Plantz et al., 2013) and multivariate analyses (Lauritsen, 2012), have found the same effects of age, marital status, race/Hispanic origin, and income.

A second similar finding is that women enrolled in college did not exhibit higher rates of victimization than non-college students (Chapter 8). Nationally, there is considerable concern over the high rates of sexual violence among college students, as revealed by recent campus climate surveys (White House Task Force Report, 2014). In analysis of the NCVS, Sinzit and Langton (2014) found that, after controlling for age, women age 18 to 24 who are not in college have slightly higher rates than those who are in college. The RSA Pilot Test found that college enrollment did not increase risk. Those currently in college had a similar victimization rate as those who were not in college.

Among several characteristics, the types of incidents that are reported on the NCVS and RSA Pilot Test surveys are similar. Both surveys indicated that these incidents infrequently involve weapons (around 8% of the time), almost always involve one male offender and involve the same percentage of offenders who are known by sight only or casual acquaintances.

There are, however, several key differences in the types of incidents measured on the two surveys. The largest difference is the percentage in which the police found out about the incident. Incidents identified on the RSA Pilot Test were three times less likely to come to the attention of the police than those on the NCVS (10% vs. 34%). This may be indicative that the incidents reported on the NCVS were more salient and were more likely to be the types of incidents that would be recalled in a crime survey than one about health and safety. For example, NCVS respondents were more likely to report injuries, emotional difficulties for at least one month, having to go to the emergency room or other hospital setting, and getting help from a victim assistance agency.

A second difference between the surveys was in the victim's relationship to the offender. The RSA Pilot Test has a higher percentage of friends or ex-friends and strangers as perpetrators, while the NCVS has a higher percentage of spouses and ex-spouses. This resulted in a higher percentage of incidents reported on the NCVS which involved intimate partners (spouses, ex-spouses, boyfriends and girlfriends). Related to this, more of the RSA Pilot Test incidents occurred at a friend's house, while more NCVS incidents occurred at the respondent's home.

There were several other differences between the two surveys, including the time of day the incident occurred and the victim's perception of the offender's use of alcohol and/or drugs.

Recommendations

This section provides the main recommendations from the study. The full list of recommendations is presented in Chapter 15.

- 1. For the ongoing NCVS, redesign the screening items that target rape and sexual assault. Expand the scope of the items to include different types of sexual contact. Expand the items on the detailed incident form to ask about the behaviors and tactics that are specific to rape and sexual assault.**

The evidence from this study suggests that one of the major reasons the NCVS rates are lower than expected is the relatively narrow focus of the NCVS screening questions. It is recommended that redesign of the ongoing NCVS (hereafter referred to as the “core NCVS”) should modify the screening questions by using behaviorally specific language to describe the types of behaviors (e.g., penetration, kissing, groping) and tactics (e.g., physical force; inability to consent) that constitute the definitions of rape and sexual assault. At a minimum, two screener questions should be used. One question should ask about unwanted sexual contact that involved force, including attempts. A second question should ask about unwanted sexual contact that happened while the person was unable to consent. However, the RSA Pilot Test suggests that more than two questions are necessary to fully enumerate all types of incidents that are within scope. The recommendation of two items is a minimum, keeping in mind that the NCVS is an omnibus survey that collects data on several other types of crimes, as well as other topics of interest (e.g., contact with the police; school crime). If additional items can be added, then it is likely that fuller measurement can be achieved.

In addition, the DIF on the NCVS does not enumerate either the specific behaviors or tactics that define rape and sexual assault. Rather, on the NCVS, a rape is operationally defined when respondents report “rape” in response to the question of how they were attacked. Besides intermingling the legal terminology of rape/sexual assault with a type of attack, this approach misses descriptors of force that make up this type of tactic (e.g., pinning down). There are no questions that ask about the different behaviors, such as different forms of penetration or unwanted kissing or groping. Finally, there are no questions related to the respondent’s ability to consent. To improve the measurement, therefore, the NCVS DIF should enumerate the type of sexual contact, including the different forms of penetration, kissing, and sexual touching and whether they were completed or attempted/threatened. Similarly, the tactics should be enumerated, including the different types of physical force that may be been used, any groping behaviors, and the ability to consent.

2. For the core NCVS, improve procedures to ensure interviews are conducted without any other persons present.

A significant percentage of the interviews on the core NCVS are administered when someone else is present in the room (Catalano, 2016). This inhibits disclosure of events that respondents may not want others to know about. The results from the RSA Pilot Test demonstrated that it should be possible to increase the privacy of the interview. This can be done in several ways. One would be to change the mode of the interview on the NCVS from interviewer-administered to

self-administered. The latter could include an ACASI or another self-administered computerized mode such as the web. Privacy can be further enhanced by training interviewers to make sure, to the extent possible, that others are not present and/or are out of vision and earshot of the interview. This should hold not just for the NCVS items on rape and sexual assault, but for all of the data collected on the survey.

About half the NCVS interviews are conducted in person and half are conducted over the telephone. Ensuring the telephone interviews are private can follow similar protocols as used on the RSA Pilot Test. Interviewers should emphasize that the respondent be in a private location without anyone else overhearing. A similar condition should be followed for in-person interviews. It is recognized that smaller housing units may not allow absolute confidentiality because it may be hard to stay out of earshot of others. If the survey is converted to self-administration (e.g., ACASI), it should still be possible to maintain confidentiality. If the survey is interviewer-administered, field representatives should be trained to maximize confidentiality as much as possible, at least ensuring physical isolation. If this is not possible, field representatives should consider accepting non-response in lieu of an interview that is not private.

3. *If redesign of the core NCVS does not fully measure rape and sexual assault, implement a separate survey within the NCVS program to collect these data.*

One goal of the RSA Pilot Test was to inform the decision on whether valid estimates of rape and sexual assault can be collected on the NCVS or whether RSA needs to be collected as part of a separate survey.⁸ The RSA Pilot Test survey did not test a revised version of the NCVS with improved measures of rape and sexual assault. Consequently, results from this study do not directly address whether a separate survey would be needed after applying the recommendations above, i.e., expanding the scope of the NCVS screening items, making changes to the detailed incident form, and improving the privacy of the interview. It will be important to test the changes to the NCVS design noted above to assess the extent to which full measurement can be achieved.

RSA Pilot Test results strongly suggest that, while it should be possible to improve the estimates on the ongoing survey, it is not likely that it can fully measure rape and sexual assault within the current constraints of the NCVS design. On the RSA Pilot Test, nine screening questions

⁸ For purposes of discussion, it is assumed that the basic design features of the NCVS will not change. For example, it is assumed that all persons 12 and over in the household are interviewed, the omnibus nature of the survey remains the same, and a significant percentage of interviews will be conducted by an interviewer.

were needed to fully enumerate the behaviors and tactics that define rape and sexual assault as well as to provide respondents with the cues to assist in recall of these incidents (table E-3). It may be difficult to fully enumerate all types of incidents by just modifying the core NCVS.

- 4. *The separate survey should include two features of the RSA Pilot Test. One is to include questions that separately ask about behaviors and tactics that define rape and sexual assault. A second is to include the RSA Pilot Test features to preserve privacy and confidentiality.***

The RSA Pilot Test did not determine the optimum number of BSQs that are needed to fully measure rape and sexual assault. On the RSA Pilot Test, 12 BSQs were used to target the specific behaviors and tactics that made up the definition of rape and sexual assault. Three of these questions asked about exposure and forcing someone to take photos/videos. These items had very low prevalence and did not contribute substantially to RSA estimates (table E-3) and could be dropped from the BSQs. When considering the BSQs for the survey, the remaining nine questions account for more than 90 percent of the incidents classified as rape and sexual assault. Six of the remaining nine questions targeted rape. These items accounted for about 83 percent of the rapes reported on the survey. The remaining three items targeted sexual assault. These accounted for 75 percent of the sexual assaults. Some items targeting rapes led to reports of sexual assaults (21% of sexual assaults) and those targeting sexual assaults led to reports of rapes (10% of rapes). These nine questions account for 93 percent of the rapes and 97 percent of the sexual assaults.

It may be possible to combine some of these questions to cover more than one type of behavior. For example, the recent Campus Climate Survey Validation Study (Krebs et al., 2016) had two multi-part questions, which resulted in prevalence rates equivalent to several other campus climate surveys (e.g., Cantor et al., 2016). On the other hand, the results of the reliability analysis for the RSA Pilot Test suggest that combining many items in a single question may lead to additional sources of variability in the estimates by making it more difficult for respondents to understand the questions. Combining questions may also reduce the number of individuals that report multiple incidents. These questions could be addressed in the final development of the separate survey.

The separate survey should include the different mechanisms to preserve the privacy and confidentiality of the interview that were implemented in the RSA Pilot Test. One of these procedures is to sample one person per household. In conjunction with the use of a graduated consent procedure, this limits the extent that others in the household know what is being asked on

the survey. In addition, the survey should use procedures to maintain the privacy of the respondent by setting up the interaction as described above in Recommendation #2.

The results of the RSA Pilot Test did not find significant differences in the incidence and prevalence rates between the ACASI and CATI interviews for the general population sample. This suggests that the selection of mode of interview for the separate survey should be based on factors other than its effect on reporting rape and sexual assault. The one significant advantage of the ACASI design is that it is more conducive to collecting details about what happened during the incident. A number of the questions on the detailed incident form relate to the circumstances that involve multiple response categories (e.g., victim-offender relationship; type of force or coercion used; reasons for not reporting to the police; type of injury suffered). For example, collecting information on the specific behaviors that occurred during the incident involves asking about different types of penetration and sexual touching, as well as distinguishing between completed and non-completed acts. The telephone survey was designed to use “yes/no” questions for each behavior to preserve confidentiality, while a self-administered survey has the advantage of presenting all of the possible options to the respondent, who can then pick the appropriate responses. As found on the RSA Pilot Test, this aspect of the design not only gives an ACASI design more flexibility with respect to questionnaire design, but it also results in a shorter interview and lower rates of missing detailed incident forms.

While potentially burdensome, the separate survey should ask respondents to provide a narrative after each incident. The results described in this report provide evidence that measurement error occurs both on the screener and the detailed incident form. With further development, it should be possible to reduce error (see below). Nevertheless, the narrative provides an important check on data quality much as it now does on the current NCVS.

5. Design the separate survey to produce incidence and prevalence estimates, as well as characteristics related to rape and sexual assault. The sample sizes should support generating these estimates on a rolling, multi-year basis.

The goals of a separate survey would be similar to the NCVS. In particular, it would be used to produce incidence and prevalence estimates of rape and sexual assault, characteristics of these incidents, and estimates of change over time. The precise sample size needed for this survey depends on a number of parameters (frequency of estimates, precision desired). There are approximately 220,000 individuals ages 12 and over who are interviewed on an annual basis for the NCVS to

produce annual estimates.⁹ Even with this large sample size, the precision is not high for estimates of rape and sexual assault. Breaking these data out by demographic groups or specific characteristics typically requires aggregating many years of data.

A separate survey that implements procedures similar to the RSA Pilot Test should produce a higher rate of rape and sexual assault victimization and would not require as large a sample size as the NCVS. For example, the RSA Pilot Test CATI survey estimated a prevalence rate for rape for women age 18 and older to be 1.9 percent with a relative standard error of 11 percent.¹⁰ If the number of interviews was increased from 5,000 to 20,000, the relative standard error would be closer to 5 percent. Estimates of rape for males would be more problematic because the prevalence rates would be expected to be much lower than for females. However, it might be possible to produce reliable estimates of sexual assault reported by males. Assuming both males and females are interviewed in equal numbers, a total of 40,000 interviews is a starting point when considering the size of the sample for the separate survey.¹¹

These calculations need to be refined, once parameters, such as the reference period and sample design, are defined more explicitly. The efficiency of the sample is dependent on features such as whether the survey is cross sectional or longitudinal, the extent of geographic and household clustering, the response rate, and the survey weighting.

6. *Investigate two methods to draw a sample for the separate survey. One is to sample from respondents to the core NCVS. The second is an independent sample.*

There are at least two possibilities for drawing the sample for the separate survey. One possibility is to administer the separate survey to the sample already participating in the core NCVS. One person in a household would be asked to participate in the separate survey after completing the NCVS. This is a design that is similar to what currently being done on the Crime Survey of England and Wales (CSEW). A second option is to draw an independent sample to administer the separate survey. Households or persons would be sampled using a general population frame, such as a list of housing units (ABS) or telephone numbers (RDD).

⁹ Note that the NCVS uses a 6-month reference period. All else being equal, it takes twice as many NCVS interviews to cover the same calendar period as the RSA Pilot Test, which used a 12-month reference period.

¹⁰This is the estimate from the RSA Pilot Test CATI survey once including all women 18 years and older.

¹¹This assumes a 12-month reference period. Sample sizes would need to be adjusted if a 6-month period was used.

One advantage of using NCVS sample is that it efficiently targets high-risk groups at very little cost. For example, it would be useful to subsample young adults, who are at highest risk of rape and sexual assault. For a survey that draws an independent sample, oversampling would significantly add to the cost of the data collection. On the RSA Pilot Test, which oversampled women ages 18 to 49, the in-person sample size was reduced by approximately 25 percent to offset the cost of screening out households. It was not efficient to conduct this subsampling for the RDD survey because of the marginal difference in the cost of screening households into the survey relative to the cost of completing the full interview once recruited into the sample (Chapter 3).

One disadvantage of sampling NCVS respondents is that it could impact participation in future waves of the NCVS. For example, if the separate survey is administered the first time the interviewer visits the household, it may deter participation on the NCVS at future waves. If this were true, then some consideration might be given to administering the survey at later waves (e.g., at the sixth or seventh interview). Another possible disadvantage of conducting the survey with NCVS respondents is the possibility that the frame of reference from the NCVS, which is focused on crime and criminal justice, may influence how respondents to the RSA survey interpret the questions on the separate survey.

One advantage of an independent sample is that it would not be constrained by the design of the ongoing NCVS and could be more flexible in its design and procedures. For example, a sample from the NCVS would need to be integrated within the rotating panel design of the survey. Participation on the separate survey may affect response rates at the next household visit, or it may influence how respondents answer the core NCVS questionnaire at the next interview.

The relative costs and response rates of the two approaches should be considered. These two parameters (cost and response rate) are directly correlated. On its face, a survey that samples from the ongoing NCVS should be less expensive than one that draws an independent sample. The costs for initial recruitment of the sample are absorbed by the ongoing NCVS. This is especially the case if groups at high risk of victimization are oversampled (see above). A recent study examining the rotating panel design of the NCVS placed the cost of completing a core NCVS interview at \$120 and \$250 by telephone and in-person, respectively (Berzofsky & Carrillo-Garcia, 2017). About half of the NCVS interviews are completed in each mode, yielding an average cost of about \$185 per complete. A separate survey that samples from the NCVS would likely be less expensive than this, since many of the interviews would be immediate follow-ups with someone completing an interview,

so no additional trips to the household will be required for these individuals. It is also possible to have the respondents complete the survey on the web, which should further reduce costs.

An independent RDD telephone survey that oversamples high-risk groups (e.g., young adults) would be more expensive than this. The RSA Pilot Test costs were more than \$200 per complete. In addition, the response rate would be considerably lower than one that samples from the ongoing NCVS. The RSA Pilot Test RDD response rate was 18 percent, although this was in metropolitan areas, which tend to have lower response rates than the nation as a whole. While the RSA Pilot Test did not find non-response bias to be an issue with the CATI survey, this conclusion is based on assumptions that could not be fully tested. In addition, the response rates for RDD continue to decline over time, which may introduce non-response bias and, at the very least, make the survey more expensive to implement. Finally, an RDD frame does not allow an easy way to ask respondents to move to the web. An independent ABS sample, which contacts respondents in-person, should be able to achieve comparable response rates to the NCVS. However, based on results from this study, the cost of conducting the survey would be approximately 4 times higher than an independent RDD survey (Chapter 13).

When designing the separate survey, an initial cost assessment should be completed to make a more concrete comparison of the two methods of sampling. Several different designs should be specified (e.g., oversampling parameters, mode of interview, reference period, and sample size) and costs generated for doing the survey as part of the core NCVS and for different types of independent samples.

- 7. Chapter 15 lists a number of recommendations on design features of the RSA Pilot Test that should be used on the separate survey, including features that should be changed: (a) limiting the number of detailed incident forms to three, (b) simplifying DIF questions on unwanted behavior, (c) expanding the list of options used on the RSA Pilot Test to measure force, (d) linking the force questions to specific behaviors, and (e) considering multiple criteria when defining an incident due to inability to consent.**

When respondents reported more than one incident, the number that refused to fill out a detailed incident form increased. Filling out multiple detailed incident forms adds to the length of the interview (Chapter 6). Some respondents became frustrated, especially CATI respondents, when asked about the details for incidents that were very similar, especially incidents involving unwanted sexual kissing or groping, which in some cases were not particularly salient. When capping incidents

at three, it will be necessary to develop an estimation procedure, using the responses to the screener, to impute for the incidents that were not collected because they were lower priority for individuals reporting multiple incidents.

The questions on the DIF that asked about unwanted behavior (D1a – D1d; D2a – D2e) were found to be problematic and were a source of error. These items had the highest levels of missing data on the interview. Behavior coding of a sample of CATI interviews found that interviewers had a hard time administering these questions. There were several instances in which the responses to these items were not consistent with the narratives and these items had relatively low reliabilities. Chapter 15 provides recommendations on changes to the wording of these items.

The item used to measure physical force should add options related to grabbing and groping. A number of respondents wrote this type of response into the “other specified” options. The omission from the list was one reason for inconsistencies observed in the re-interview.

The RSA Pilot Test had separate questions on behaviors and tactics. A respondent could report several unwanted behaviors, as well as several tactics, but there was no linkage between the two. It is recommended that the behavior and tactic questions be interwoven so there is a direct linkage.

Classification of the incident as due to being unable to consent was based on the response to one of three questions: (1) a direct question on ability to consent (question G12a), (2) the respondent was passed out for part of the incident (question G10), or (3) the respondent was passed out for all of the incident (G10). The legal criteria defining alcohol/drug-related incidents of rape and sexual assault is that the victim is unable to consent and that the perpetrator knows this. If the respondent is unconscious or asleep, this clearly meets this standard. The standard for being too intoxicated to make decisions, but still being conscious, is more difficult to prove in a court of law and, by definition, more difficult to measure on a survey. For this reason, the analytic classification scheme used the respondent’s self-report that she was unable to consent. Overall, the evidence cited in the report indicates that most who said they were unable to consent met some criteria of being intoxicated to the point that they had trouble making decisions. This evidence also provided some indication that another individual would recognize the victim was not in condition to consent. Of course, the indicator is subject to some error with both false positives and false negatives, and the phrase “unable to consent” is open to some interpretation and likely contributes to this error. We recommend including the same, or related, indicators of drug and alcohol use that were measured on

RSA Pilot Test on the supplemental survey to give analysts the ability to see if results differ by how inability to consent is defined.

Limitations and Additional Research

The RSA Pilot Test has several limitations. One is that the sample was restricted in several ways. Males were excluded from the sample (Chapter 2), so the survey language needs to be adapted and tested with men. The ACASI survey was not administered to anyone over the age of 49. Neither the ACASI or CATI survey was administered to anyone between the ages of 12 and 17. The survey was conducted only in five large CBSAs. While there is no reason to believe there are unique effects of the methodologies implemented on RSA Pilot Test for men, older persons, or residents of these particular cities, future development work should expand to a larger universe.

Another limitation is that the overall response rate was below what could be achieved on the NCVS. The analysis did not find significant nonresponse bias when comparing the 18 percent response rate for the CATI and the 40 percent response rate for the ACASI. Nonetheless, the effects of nonresponse on the estimates cannot be ruled out when a survey with a higher response rate is implemented.

As noted, several questions should be addressed as a new set of measures of rape and sexual assault are incorporated into the NCVS program. A high priority is to revise the methods used on the core NCVS and assess how much this improves the measures. The NCVS is currently undergoing a redesign of the entire instrument, so the results of this study should be considered when making changes to the items related to sexual assault. A second priority should be to assess the two alternatives proposed for administering a separate survey (sampling from ongoing NCVS sample vs. drawing an independent sample). Cost data from the ongoing NCVS, as well as cost estimates to conduct a survey using an independent sample, should be used to map out these differences in more detail.

It will be important to examine the recommended procedures within the context of the NCVS framework and Census Bureau operations when finalizing the methodology for measuring sexual victimization. As noted above, there are a number of differences between the rotating panel design of the core NCVS and those used on the RSA Pilot Test. If prior methodological studies are any indication (e.g., Biderman et al., 1986), these differences can have a significant effect on the incidence and prevalence rates produced by the ongoing NCVS.

1. Background and Purpose

Over the past 2 decades, there have been a number of competing national estimates of the level of rape and sexual assault in the United States. The official estimates of these crimes based on the National Crime Victimization Survey (NCVS) have typically been lower than estimates obtained from surveys contracted by other Federal agencies and by private groups (Black et al., 2011; Koss & Gidycz, 1985; Tjaden & Thoennes, 2000; Kilpatrick, 2007; Fisher et al. 2000). For example, estimates of rape from the National Intimate Partner and Sexual Violence Survey (NISVS) are approximately four times higher than comparable NCVS estimates (Tjaden & Thoennes, 1998). Estimates of rape from the NISVS (Black et al., 2011) are more than 10 times higher than the NCVS (Truman, 2011).¹² These differences have resulted in debate over the best method for collecting self-report data on rape and sexual assault (Fisher & Cullen, 2000) and, perhaps more importantly, has led to some confusion on the level of rape and sexual assault in the nation (e.g., Gilbert, 1997; Lynch, 1996; Rand & Rennison, 2005; Bialik, 2013).

The NCVS is an omnibus survey that is designed to provide data on trends in victimization for persons age 12 or older that are independent of police reporting. It produces national rates for a wide range of predatory crimes (e.g., robbery, rape, assault, burglary, theft), as well as emerging crimes such as identity theft, fraud and stalking. The goals of the survey also include providing data on the characteristics and consequences of criminal victimization, as well as information on issues related to crime and victimization, such as public perceptions of crime and safety. This is in contrast to more specialized surveys that concentrate on data on sexual violence.

Some of the differences in the estimates noted above result from different definitions of rape and sexual assault (GAO, 2016). The NCVS, for example, emphasizes forcible rape, while other surveys include tactics other than direct physical force (e.g., inability to consent due to drugs and alcohol). Even when surveys use comparable definitions, there are still large differences in the estimates. This is due to the different methodologies used to elicit reports of these events. Features such as the sample design (e.g., rotating panel vs. cross sectional), mode of interview (in-person, telephone, self-administered), response rate, context (crime survey vs. health survey), interview

¹²The 2010 National Intimate Partner and Sexual Violence Survey (NISVS) prevalence rate for rape is 1.1 per 100 adult females. The NCVS incidence rate for rape and sexual assault 0.13 per 100 females 12 and over. This comparison underestimates the differences between the two surveys because the NISVS is a prevalence estimate and only counts a victim once, regardless of how many times she was victimized. The NCVS is an incidence rate, which counts all incidents that happened to each victim. In addition, the NCVS includes sexual assaults, while the NISVS estimate does not.

setting (in private setting vs. others present) and length of reference period (lifetime, 12 months, 6 months) are all design features that can have significant impact on the number of victimizations reported (e.g., Lynch, 1996).

One of the most obvious differences are the methods to screen and count victimizations between the NCVS and many of the other surveys. The NCVS uses a two-stage design, which initially screens for victimizations and follows up with a detailed incident form (DIF) for each incident reported from the screener. The incident is classified into a particular type of crime based on responses to the DIF, which include questions related to the criteria related to defining the incident as a crime. Many of the other surveys are one-stage designs, which use responses to the screening questions to count and classify victimizations.

The differences in methodologies, in part, stem from the rationale for each type of survey. As noted above, the NCVS has a primary goal of generating annual estimates of incidence and prevalence of different types of criminal events. The reason NCVS uses the DIF to classify the event is to determine if the incident meets the definition of a particular type of crime and to count the number of incidents for an annual period. Other surveys have focused on the harm and consequences associated with sexual violence and their focus is less on annual estimates and more on events that have occurred over the victim's lifetime.

BJS initiated the Rape and Sexual Assault Pilot Test (RSA Pilot Test) to develop recommendations on the best methods to collect data on rape and sexual assault within the NCVS program. The project had three objectives:

1. to develop and pilot test a design based on best practices to collect self-report data on rape and sexual assault using an in-person, audio computer-assisted self-interview (ACASI) questionnaire
2. to develop and pilot test a comparison design using random digit dialing (RDD) and a computer-assisted telephone interview (CATI) survey
3. to conduct detailed analytical comparisons of the two designs against each other and the existing NCVS program.

One goal of the project was to improve the data collection methodology and measurement within the NCVS program. The RSA Pilot Test developed recommendations for the NCVS program by testing several different designs that implemented features thought to improve data quality relative to the NCVS. As part of this goal, the project was to assess whether it was possible

to accommodate improvements in measurement within the NCVS core instrument or whether a separate survey was needed.

A second goal was to contribute to knowledge on the best methods to collect data on rape and sexual assault. There have been several reviews of methods used to conduct victimization surveys collecting rape and sexual assault data, including why estimates differ between the NCVS and other surveys (Lynch, 1986; Cook et al., 2011). Five issues were investigated:

1. How and why do estimates from a survey using behaviorally specific questions differ from the NCVS?
2. Is there a difference in estimates of rape and sexual assault between self-administered and interviewer-administered modes?
3. Are there significant effects of non-response bias for a survey on rape and sexual assault for either the in-person ACASI and RDD CATI surveys?
4. What are the advantages and disadvantages of a one-stage vs. two-stage design?
5. What is the data quality (validity and reliability) of estimates from a questionnaire using behaviorally specific questions (BSQs)?

1.1 Critical Design Components for Rape and Sexual Assault Surveys

After reviewing the literature related to the measurement of rape and sexual assault, as well as considering the goals and objectives of the NCVS program, four design features were considered when designing the two pilot surveys:

- maximize privacy and confidentiality
- maximize population coverage and minimize nonresponse bias
- minimize measurement error associated with the questionnaire
- minimize emotional trauma.

These are considered critical when trying to collect unbiased estimates of rape and sexual assault in a general population survey (Koss, 1993; Kruttschnitt et al., 2014).

1.1.1 Mode of Interview and Maximizing Privacy and Confidentiality

When reporting on sensitive behaviors, such as rape and sexual assault, the privacy and confidentiality of the interview is an important element to eliciting reports by a respondent. The survey literature defines several different types of sensitive survey questions (Tourangeau & Yan, 2007). One type concerns the intrusiveness of the question content. A question is intrusive when it invades a respondent's privacy and is considered inappropriate when asked in normal conversations. Asking about rape and sexual assault would be considered intrusive. A second type of sensitive question is when the answer might disclose the information to a third party. This disclosure may result in negative social or physical consequences. For instance, a respondent may feel threatened by a spouse or intimate partner when reporting an incident on the survey. Even if the other person was not the perpetrator, revealing the information to a third party may be perceived as stigmatizing (Koss, 1993).

A survey can be administered by an interviewer or a survey can be completed by respondents on their own. Methodological research has shown consistently that self-administered modes produce more reports of sensitive behaviors than interviewer-administered modes (Tourangeau & Yan, 2007), although there has not been a great deal of research on the effects of self-administration for reports of sexual violence (for exceptions see Laaksonen & Heiskanen, 2014; Cantor & Williams, 2013).

Many of the surveys on rape and sexual assault have been conducted over the telephone.¹³ These surveys have used procedures to maintain confidentiality of the survey topic, as well as the specific survey questions. To conceal the topic of the survey, the initial contact with the household presents the study as something more general than sexual violence, such as on “health and safety.” Once contact has been made with the person selected for the survey, the more specific topic of the survey is revealed. To maintain privacy while the interview is going on, the interviewer emphasizes that no one else in the household should be listening to the interview. Questions are phrased to only require a yes or no answer—this prevents someone else who might be listening from deciphering the survey topic. Respondents are also told they can hang up the telephone if, at any point, they feel they are in any danger of being overheard or don’t feel safe.

The above precautions seem to be effective judging by the relatively high rates of rape and sexual assault that are measured by the telephone surveys, although there have been very few

¹³One exception to this is campus climate surveys, which have also been conducted over the internet.

comparisons between different interview modes. For example, high rates of sexual violence on these surveys could be due to non-response bias or other features that lead to higher rates (e.g., external telescoping; Lynch, 1996). More generally, it is unclear how a self-administered survey performs relative to a telephone interview, which takes the above precautions. One hypothesis is that rape and sexual assault questions are sensitive because of fear that the information may become known to a third party. By taking precautions to prevent this type of disclosure, a telephone interview may produce comparable rates to a self-administered survey.

An in-person interview poses a similar set of threats to confidentiality and privacy, except the channel of communication is a bit wider than the telephone. There is both verbal and non-verbal communication that occurs. This could foster rapport and ensure confidentiality or have the opposite effect on feelings of privacy and confidentiality. There is also a different set of challenges to keeping a third party away from the survey interview. Without special precautions, it may be difficult to isolate the respondent so that no one is able to overhear the interview. NCVS interviews that are done in person have a significant number that are conducted with a third party present (Catalano, 2016). Those that are conducted with a third party present have significantly lower rates of rape and sexual assault than those conducted when no one else is present (Coker & Stasny, 1994).

In order to preserve the confidentiality and privacy of the interview, one of the RSA Pilot Test surveys used a self-administered mode of interview because of the advantages of using self-administration when asking about sensitive topics. In particular, audio computer-assisted self-interview (ACASI) was used in conjunction with an in-person visit to the house (see discussion below). This mode offered the ability to keep the interview private from both the interviewer as well as any other third party that might be present when the interview was being completed. The second RSA Pilot Test survey was conducted by telephone, using the graduated consent procedure described above. This allowed the study to assess the relative advantages of each of these approaches.

1.1.2 Minimize Coverage Error and Nonresponse Bias

Coverage error refers to the extent to which the sampling frame includes the target population. The NCVS is an area sample, based on enumerating housing units and individuals within those units. The more specialized rape and sexual assault surveys have been random digit dial (RDD) surveys. Generally speaking, area samples have better coverage than RDD surveys. The address frames that are used for area samples do have undercoverage, especially for highly mobile

populations, but they are considered to be the best that can be achieved when doing a survey (Groves et al., 2004). Telephone surveys have lower rates of coverage. There is a small portion of the population that does not have any telephone service (2.5%; Blumberg & Luke, 2014). The list-assisted methodology that is used to generate the RDD frame also excludes a portion of the landline telephones, although exactly how much is open to debate. Recent discussions have found it to range from 5 percent to 20 percent (Fahimi et al., 2009; Boyle et al., 2009).

Of more concern for the NCVS program is the possible bias due to nonresponse. Area samples using an in-person contact typically have much higher response rates than telephone surveys using an RDD frame. Response rates for all types of surveys have been dropping over the last 10 years. This drop has been especially dramatic for RDD telephone surveys (Curtin et al., 2005; Brick & Williams, 2013). For example, in 1992 the National Women's Study (Kilpatrick et al., 1992) reported a response rate of 85 percent, the National Violence Against Women survey (Tjaden & Thoennes, 2000) conducted in 1996 had a household participation rate of 72 percent, and in 2010 the NISVS had a response rate between 27 percent to 33 percent (Black et al., 2011). The steady decline in response rates for RDD surveys has led several national survey programs to abandon the RDD methodology (National Household Education Survey, Montaquila et al., 2013; Health Information National Trends Survey, 2009).

Response rates for in-person surveys have also been declining over this same time period, although to a lesser degree than telephone surveys. For example, the household response rate for the NCVS has declined from 92 percent in 2000 to 82 percent in 2014. Nonetheless, even the lowest response rate for the NCVS is considerably higher than the response rate for any RDD survey.

Logically, as nonresponse rates go up, the greater the chance the respondents may differ from the nonrespondents in a way that poses significant risks for biased estimates. For example, one might believe that those who are least concerned about the risks of victimization may also be the least likely to agree to participate on the survey. If this was true, survey estimates may be too high because the sample overrepresents the victims. Empirically, however, a number of studies have not found the response rate to be strongly related to bias across a wide range of topics (Keeter et al. 2000; Keeter et al. 2006; Merkle & Edleman, 2002; Groves, 2006). There has not been a great deal of work in this area for surveys measuring rape and sexual assault. Recent work investigating campus sexual assault (Cantor et al., 2016; Krebs et al., 2016) found some evidence of a small positive bias. That is, surveys with low response rates had slightly higher rates of victimization than those with higher response rates. This research suggests that the bias seems to level off around a response rate of 30 percent, at least for student surveys. Consistent with this conclusion, in a pilot test associated

with the NISVS, Peytchev et al. (2009) found that when moving from a response rate around 20 percent to around 30 percent, the estimates for certain types of sexual violence changed, but the change in the estimates was relatively small.

Given the importance response rate has played in evaluating survey quality, the pilot surveys varied the methods of initial contact. To maximize the response rate, an in-person visit was paired with an ACASI mode to leverage the advantages of self-administration (see discussion above). An in-person visit capitalizes on both the better coverage and response rate properties associated with this mode of contact. The telephone interview was based on an RDD frame. The study then compared the relative advantages and disadvantages of each approach for the NCVS program.

1.1.3 Minimize Measurement Error Associated with the Questionnaire

Measurement error refers to the extent that the responses to the survey deviate from the “true value” (Groves et al., 2004). Any design feature can contribute to measurement error, including the mode of interview (see discussion above). The design of the questionnaire has been a particular focus as a source of error for the measurement of rape and sexual assault (Lynch, 1996; Kilpatrick, 2007; Fisher & Cullen, 2000). The measurement of rape and sexual assault poses unique challenges not only because of the sensitive nature of the topic but also because of the ambiguities associated with defining eligible events. The terms “rape” and “sexual assault” are not uniformly understood by the general population. There is a common image that a rape involves a stranger physically forcing a woman to have vaginal sex in a dark alley. In fact, studies find that most rapes involve individuals who know each other. In addition, legal statutes include all types of sexual penetration (vaginal, oral, or anal) of a man or a woman, by objects besides a penis (digital or with other objects) and do not necessarily involve physical force (alcohol or drug facilitated) (Tracy et al., 2013).

To measure rape and sexual assault, the best practice is to use behaviorally specific questions (BSQ), which enumerate the behaviors, tactics, and absence of consent (Kruttschnitt et al., 2014). The questions offer different combinations of behaviors and tactics that make up the definition of rape and sexual assault. For example, on the National Violence Against Women survey, the items that cover completed rape were –

Has a man or boy ever made or tried to make you have sex by using force or threatening to harm you or someone close to you? Just so there is no mistake, by sex we mean putting a penis in your vagina?

Has anyone, male or female, ever made or tried to make you have oral sex by using force or threat of harm. Just so there is no mistake, by oral sex we mean that a man or boy put his penis in your mouth, or someone, male or female, penetrated your vagina or anus with their mouth or tongue?

Has anyone ever made or tried to make you have anal sex by using force or threat of harm? Just so there is no mistake, by anal sex we mean that a man or boy put his penis in your anus?

Has anyone, male or female, ever put fingers or objects in your vagina or anus against your will by using force or threats?

Each question specifies the behavior (e.g., “putting his penis in your vagina”; “penetrated your vagina with their mouth or tongue”) and the tactic (“make you...by using force or threatening to harm you or someone close to you”).

Using very specific language is considered better than using more colloquial words such as “sexual relations” or “sexual contact” because some may not think of the encounter in those terms or the event may not be stored in memory under this rubric. For example, digital penetration may not be recalled when asking about “sexual contact.” The sheer number of questions that are used to screen for the particular behaviors on the sexual assault surveys also give the respondent more time to recall events that may be of interest. Fisher (2009) compared the smaller number of screening items used on the NCVS, which are far less specific and use terms like rape and sexual assault, to a series of screening items using BSQs. She found the rates of rape measured on the survey using the BSQs were about 10 times higher than those captured by the NCVS screening items.

One issue associated with BSQs is their complexity. By asking about both behaviors and tactics in the same question, respondents have to consider multiple concepts when retrieving memories and judging whether they qualify as asked by the question (Cook et al., 2011). A respondent who has been forcibly kissed or groped, for example, may respond to the penetration item because she remembers being forced to have some type of sexual contact. Or a respondent may report an unwanted sexual experience even if does not involve physical force. Respondents’ ability to retain in short-term memory all of the conditions associated with the question may lead to error. This can be problematic if the screening items are used to classify the event in the analysis, as most of the sexual assault surveys do.

A second criticism of this approach is, in the attempt to reduce underreporting, BSQs can lead to a much larger number of false positives. Fisher and Cullen (2000) sum up this criticism:

The critics accuse feminists of merely finding what they set out to find—of using research methods that are so flawed and ideologically biased that they present estimates of rape that are inflated many times over. . . . In particular, critics contend, the definitions of rape and the survey questions used to measure rape are so broadly or poorly phrased that they “pick up” and count as rape a wide diversity of conduct, most of which could hardly be considered criminal in a legal sense. (p. 320)

Critics of the BSQ approach point to the large proportion of respondents who are classified as a rape victim on surveys using this approach but do not label it as such (Gilbert, 1997). For example, Koss (1988) found that only 27 percent of college students who reported an event that met the legal definition of rape on the survey identified it as such. Critics say that college-educated women should know they have been raped. The rebuttal to this criticism is that most individuals do not know what constitutes the legal elements of a rape, especially if it involves an intimate partner or a friend. Note that even the legal definition of rape has changed significantly over the last 50 years (Fisher & Cullen, 2000). It is only in the last few years that the FBI expanded their definition of rape in the *Uniform Crime Reports* to include elements of “non-consent.”

Similarly, laws are still evolving related to the role of alcohol and drugs and defining the incident as rape. Most (if not all) states make it a crime if the victim is unconscious or asleep because of substance use. Even when conscious, if the victim is so intoxicated that she can't make a decision about consent, it is considered a crime. The condition of the perpetrator is not relevant, nor is whether the alcohol was consumed voluntarily. It isn't clear how many individuals are aware of this when defining incidents as a crime.

Finally, one coping mechanism of victims of sexual violence is to deny that it occurred. This would make it less likely for victims to think of the event as a rape or sexual assault.

It is not clear how often false negatives and false positives occur. One method to explore this possibility is to collect qualitative information about the circumstances of particular incidents to get a better sense of the context of the event. Koss and Gidycz (1985) conducted a detailed follow-up interview with 68 college students who reported a rape by marking the appropriate BSQ questions. According to their analysis, only three of the 68 individuals would be considered a false positive (i.e., not be classified as a victim of a rape according to legal statutes).

Testa et al. (2004) did a similar exercise for a sample of 1,000 women selected using an RDD sampling frame. A survey using BSQs was administered using an in-person interview, and each respondent was asked to describe the incident. The description was then classified as being one of the types of events described by the BSQ. For those events classified as rape as part of the BSQ instrument, between 81 percent and 94 percent were determined to be correct based on the follow-up interview. Two other interesting patterns emerged from this study. One was that the reports of attempted rape from the BSQs did not have strong agreement with the coding of the narratives. Only about 40 percent of the BSQ reports that were an attempted rape were coded as such. Second, there were a significant number of narratives that were classified as a rape by the coders but stemmed from BSQs that were targeting something other than rape. Among the narratives the coders classified as a rape, 20 percent came from BSQs that were not intended for this type of event. One limitation of this study was there were only 16 incidents identified from the BSQ questions as a rape. A second limitation is that respondents were primed on which questions they answered before being asked the narrative questions. This may have introduced a consistency bias on the part of the respondent (Cook et al., 2011).

A second method that can be used to identify false positives on a larger scale is to implement a two-stage survey design like that used on the NCVS. The NCVS differs from surveys on rape and sexual assault by following up each affirmative response to the initial victimization screener questions with a series of follow-up questions that collect the details about what happened, including the month of occurrence, the specific behaviors that occurred, the type of force, and other details about the incident that are of interest. The attribute information from the DIF is used to determine whether the incident meets legal standards related to the crime classified as a rape or sexual assault. For example, if someone reports a sexual assault on the screener, the DIF includes questions about the type of behavior and tactic that was used as part of the incident. An advantage of a DIF is that it makes it less critical for the respondent to classify their incident solely based on the screener item. As noted above, the complex nature of these questions may lead to respondents to report incidents that may be assaults but not the specific type of incident the question is asking about. Furthermore, respondents may think of events after the relevant screening item is administered.

For reports of rape and sexual assault, a DIF could be used to collect specific information about what happened with respect to the elements that define the event as a crime, such as the specific behaviors and tactics that were used. This is a significant advantage for the NCVS, given the goals of counting criminal events. Fisher (2009) used a two-stage design with BSQs as part of the victimization screener and a DIF to classify the incident. This study found that both false positives and false negatives occur if the BSQ is used to classify the incident. About three-quarters of the

incidents that came from the BSQs designed to measure rape were classified as something else (some not being a crime at all). These would be considered false positives. However, about half of the incidents that screened in as something other than rape were eventually classified as a rape. These would be considered false negatives. The end result was that the overall estimate of rape was about the same as estimated using the BSQs, but the item which generated the report was different.

This result suggests that measurement error associated with the BSQs is subject to both false positive and false negative errors. Some respondents report incidents that don't meet the requirements for rape, while others are reporting rapes that are not classified as such by the BSQ, perhaps because their memory has been jogged at that point or they are more comfortable reporting it at a different point of the interview. This study does not conclude that the false positives overwhelm the false negatives, as suggested by critics of the BSQ methodology.

The disadvantage of the two-stage method is that the DIF is also subject to measurement error. If respondents do not interpret the questions as intended, this will lead to erroneous classification. For example, about 15 percent of the NCVS DIFs are revised in a post-editing process because the narrative information collected by the interviewer does not match what was filled out on the DIF. Cook et al. (2011) also make this point by hypothesizing several issues with the follow-up questions used by Fisher.¹⁴ One was that the item asking about physical force did not enumerate what might have been meant by this term and that respondents may have used too narrow a definition. A second was that some respondents may have viewed the wording of some of the DIF items to "... sound similar to traumatizing reactions the woman may have heard from others when relating her experience, akin to secondary victimization...."(p. 210). Both points need to be considered when using a detailed incident form. In particular, follow-up questions are subject to measurement error, as well as the idea that items on the DIF may be viewed as redundant or even questioning what was said in response to the behaviorally specific screener items. Nonetheless, this study suggests problems with the one-stage approach as the only mechanism to classify the event. For the NCVS, which has as its primary mission estimating the incidence and prevalence of crime, this is an important issue.

¹⁴Not summarized below are two criticisms made by Cook et al. (2011), which are not correct. First, they criticize Fisher as using a "criminal justice" definition of rape by restricting it to just penile-vaginal penetration. This is not correct. Fisher's BSQ screener included all types of penetration that are included in current legal definitions. A second incorrect critique was to claim that Fisher's criteria for defining rape was that respondents had to indicate both completed and threatened physical force. The criteria actually used was that either completed or threatened force had to be marked.

Both of the RSA Pilot Test surveys used BSQs to screen for incidents, as this is considered best practice. Since the NCVS is used to classify and count incidents of crime, both surveys administered a DIF to classify the incidents as a crime or not, as well as the particular type of sexual violence (e.g., rape, sexual assault, unwanted sexual contact). The study compared the classification of incidents according to the BSQs to that suggested by the DIF to evaluate data quality.

1.1.4 Minimize Emotional Trauma

Asking about sexual violence is likely to bring up negative emotional feelings for some respondents. The NCVS currently provides respondents with hotline numbers they can call. However, there is not any specific training of interviewers to identify respondents who might be upset during the interview. Of course, the NCVS is a general crime survey, with the rape and sexual assault questions being only a small subset of incidents that are reported by respondents. If the NCVS program were to conduct a more specialized survey on rape and sexual assault, it would be important to use procedures that are sensitive to respondents who become emotionally distraught during the interview.

There have been a number of sexual violence surveys that have successfully implemented procedures to identify victims who become emotionally upset, as well as provide respondents with resources to work through these issues. These studies have shown that asking about traumatic events may evoke negative feelings, but they are at a level that can be managed. This research also shows that these same respondents express very positive feelings through participation in the survey and typically rate their participation as being more positive than those who do not report any victimization.

There are several procedures that need to be implemented to maintain confidentiality, as well as being sensitive to the emotional and physical well-being of the respondent. One is to maintain the privacy of the interview. Interviews should be done without anyone else knowing about the survey questions or the answers that are given. The graduated consent procedure, used by NISVS and other sexual violence studies, has been found to be successful when recruiting and interviewing sample members.

The procedures to maintain privacy when conducting the interview vary somewhat by interview mode. For the telephone interview, the interviewer should make sure the interview is done without anyone else overhearing the conversation. This should be emphasized to the respondent

during the interview. A second procedure that has been used is to structure all of the questions so only a “yes” or “no” response is required. This prevents anyone within earshot from knowing what the conversation is about.

For an in-person interview, the environmental factors associated with maintaining privacy are different. A third party can potentially hear both the respondent and the interviewer. This threatens the graduated consent procedure, since the statement of the purpose of the study occurs when the informed consent is administered to the selected respondent. If a third party can overhear the conversation, even if in a different room, it is problematic. Finding an area to conduct the interview requires negotiating a private space within a household, which may have both physical and social impediments. For example, household members may become suspicious if the interviewer insists the survey be conducted in a private room. Administering both the informed consent and the survey with the ACASI allows the interviewer to control the interviewing environment for an in-person visit. Even if someone can overhear the conversation between the respondent and interviewer (e.g., in another room), the sensitive information is communicated by self-administration.

A second important component to administering a survey on sexual violence is to design procedures to be sensitive to the feelings of the respondent. The consent and introduction to the survey should inform respondents that the information being requested is sensitive and may evoke strong emotions. As noted above, the administration of the detailed incident form has been hypothesized as raising issues related to appearing to question the veracity of the respondent’s first response on the screener. Appropriate language recognizing this at the beginning, as well as at key items, should be sensitive to this.

A third component is to train interviewers to recognize when the respondent may be getting upset and checking in with the respondent to make sure they are able to continue the interview. These trainings might differ somewhat by mode of interview because of the channel of communication unique to each mode (verbal vs. visual and verbal).

The RAS Pilot surveys each implemented procedures to minimize emotional trauma. As noted above, these procedures may differ somewhat between an in-person visit using ACASI and a telephone interview. However, both surveys included methods to evaluate how successful the procedures were by debriefing the respondents and asking interviewers to observe how respondents reacted to the survey, as well as the interviewing environment.

1.2 Overview of this Report

The remainder of this report describes the results of the field test that compared the two RSA Pilot Test surveys (ACASI and telephone), as well as to the ongoing NCVS. Chapter 2 summarizes the design of the two surveys, as well as the activities used to develop each of these surveys. Chapter 3 describes the sample designs for the two general population samples. The in-person design used address-based sampling (ABS), while the telephone survey used a dual frame RDD, which sampled both landline and cell phones. This chapter also describes the process to develop the survey weights for each sample.

Chapter 4 describes the data collection procedures. It summarizes how the field operations were completed from interviewer training to the collection of the data. This chapter also describes the recruitment of two auxiliary samples. One was recruited by asking for volunteers age 18-29 from a Craigslist ad (volunteer, or VO, sample). These volunteers were randomly assigned to either the ACASI or CATI survey. The VO sample provided supplementary information on individuals who were expected to report victimizations at a higher rate. The VO sample is not combined with the general population data but supplements the analyses that compare the ACASI and CATI results. The results for the VO sample are primarily provided in the appendices. There are two exceptions to this. First, when there are significant differences between the general population and VO sample, the data are provided in the main report. Second, the two samples are combined in the analysis of the re-interview results for the DIF (Chapter 12). The second auxiliary sample was of known victims recruited from service provider organizations such as rape crisis centers. These participants provided information on how a group of victims who had been previously subject to traumatic experiences reacted to the survey. The results are provided in Appendix K.

Chapter 5 presents the response rates for the ABS and RDD sample. Chapter 6 describes two aspects of the surveys – the amount of time it took to fill out each survey and the characteristics of the respondents to each survey. Both the demographics of respondents, as well as other characteristics collected on the survey (e.g., income, tenure, marital status), are presented by type of sample and mode of interview. The characteristics for both the general population and VO samples are provided so the reader has an idea of how the samples compare, as well as how they may differ by mode.

Chapter 7 presents the incidence and prevalence rates for the two surveys (ACASI and CATI). This chapter begins by describing the classification algorithm used with the DIF when

determining the type of crime. This process included examining the narratives provided by respondents to assess the performance of the DIF. Sections in this chapter also provide the prevalence and incidence rates for the two surveys and compare the Pilot Test results to the NCVS and the NISVS.

Chapter 8 describes the characteristics of the victims and the incidents reported on each of the two surveys. The initial section of this chapter describes the types of individuals who are most likely to be victimized and makes comparisons between the ACASI and CATI interviews. The next section uses the DIF to describe the types of incidents that were reported in the study. The DIF included a number of items focusing on characteristics related to the behaviors, tactics, and context of the incident. For example, questions were asked about the type of physical force used, measures of alcohol/drug use, and actions victims took when the incident was taking place. These data are presented as one set of criteria to assess whether events should be included as a rape or sexual assault. This chapter also includes a section comparing the RSA Pilot Test to national NCVS data for women ages 18 to 49 for key measures of victimization.

Chapter 9 provides a more detailed analysis of key features of the BSQs and the DIF. The initial section examines the incidents that were determined to be eligible from the BSQs and the extent to which duplicate incidents were reported across different BSQs. The next section examines measures of recall and external telescoping. The next few sections describe the missing data at different levels, including from the BSQs, the willingness to fill out the DIF, and item nonresponse for specific items on the DIF. The final section does a detailed comparison of the classification of items using the BSQ and the DIF. The narratives are used to assess the quality of each source of data for completed and attempted rapes.

Chapter 10 describes several analyses assessing nonresponse bias in the ACASI and CATI samples. As noted above, one drawback of the CATI is that the response rate is expected to be significantly lower than the ACASI. This chapter explores whether this difference has implications for the quality of the estimates for rape and sexual assault.

Chapter 11 examines several different criteria to evaluate the quality of the data on rape and sexual assault. The first section describes analysis of a series of vignettes that were administered to respondents at the end of the interview. These vignettes experimentally varied situations and contexts to assess the performance of the survey questions on physical force, alcohol use, and coercion. The next section examines the performance of the event history calendar that was used on

the ACASI. The final two sections examine the performance of the CATI survey by analyzing the interactions between the respondent and the interviewer, as well as estimating interviewer variance.

Chapter 12 reports the results of analysis of a re-interview that was conducted. A subset of respondents who reported a victimization were selected to be re-interviewed approximately 2 weeks after the first interview. This chapter describes measures of reliability for each of the samples at the screener level for the BSQ items and key items on the DIF.

Chapter 13 describes the relative costs of the ACASI and CATI interviews. Relative costs are also described for the landline and cell phone interviews conducted as part of the CATI survey. Chapter 14 describes an assessment of key human subject issues. After the survey was completed, respondents were asked a series of questions asking them about their experience with taking the survey. This included their emotional reactions and whether they would be willing to take the survey again. There was also a series of items asking the interviewer to report on their observations about the survey. For the ACASI, interviewers were asked to assess any difficulties with finding a private space in the household, whether there was a third party in the room when the survey was being taken, and an assessment of the respondent's emotional reactions to the survey. A similar set of observations were asked of the CATI interviewers.

The final chapter summarizes the results and makes recommendations for integrating a survey on rape and sexual assault into the NCVS program, as well as changes to the procedures and questions that were used on the RSA Pilot Test.

2. Development of the Surveys

Chapter 1 outlined the basic elements that are of concern when administering a victimization survey collecting data on rape and sexual assault. This chapter provides the specifics of how the RSA Pilot Test ACASI and CATI surveys were developed and administered. In the first section, an overview of the sample design is provided. The second section describes the survey procedures used to carry out each survey. The third section covers the design of the victimization screener and the Detailed Incident Form (DIF). The fourth section describes the structure of the interviews. The fifth section describes the procedures used to test and finalize the questionnaire.

2.1 Overview of Sample Designs

The study targeted five core-based statistical areas (CBSAs) – New York City, Los Angeles, Dallas, Phoenix, and Miami.¹⁵ For each area, general population samples were drawn for both in-person and telephone interviewing, volunteers were recruited from Craigslist, and clients of rape crisis centers were recruited.

2.1.1 General Population Sample

A general population household sample was drawn for an in-person survey and a telephone survey. Chapter 3 provides the specifics of how the sample was drawn for each mode. In this section, we provide a general overview and rationale for the designs.

In-Person Contact (ACASI)

A household sample was drawn from the five CBSAs. The sample was restricted to adult females age 18-49 years old. The Bureau of Justice Statistics (BJS) intends to ultimately collect data on rape and sexual assault for both sexes and for ages 12 or older (as is currently the case for the National Crime Victimization Survey [NCVS]). The sample was restricted to women because the

¹⁵New York–Newark–Jersey City, NY–NJ–PA Core-Based Statistical Area, Los Angeles–Long Beach–Anaheim, CA Core-Based Statistical Area, Dallas–Fort Worth–Arlington, TX Core-Based Statistical Area, Phoenix–Mesa–Scottsdale, AZ Core-Based Statistical Area, Miami–Fort Lauderdale–West Palm Beach, FL Core-Based Statistical Area.

prevalence of rape and sexual assault among men in the general population is at least 10 times lower than for women (Black et al., 2011). Neither the NCVS nor the NISVS can produce a reliable 12-month estimate for males. The NISVS, for example, interviewed fewer than 30 male victims out of the 10,000 who completed the survey. The NCVS conducts approximately 100,000 interviews with males in a year and also cannot produce a reliable annual estimate for rape and sexual assault for this group. Because of this low prevalence, including males was deemed impractical.

For this reason, this project was developed for use with female respondents only. Prior studies including males and females have used identical, or slightly adapted, questions and interviewing protocols. Findings from these studies do not reveal special measurement issues associated with the respondent's sex. For example, the NISVS interviews both males and females with very similar protocols and questions, as do most of the surveys on campus sexual assault (e.g., Krebs et al., 2016; Cantor et al., 2016). The most significant differences are adapting the behaviorally specific questions (BSQs) to reference appropriate anatomical characteristics of a respondent. As the NCVS program moves forward with finalizing a design for collecting rape and sexual assault data, the methodology can be adapted and tested with males. Individuals under 18 years of age were not included for cost reasons. The amount of effort required to obtain parental consent to complete interviews, as well as develop separate protocols for respondents of this age, would have added significant costs to the project.

The sample was restricted to the 18-49 age group primarily for reasons of cost. Women over age 49 are at very low risk of rape and sexual assault, and concentrating on the younger age groups made the sample design more efficient for purposes of evaluating the methodologies (see Chapter 3 on the sample design). Those in the 16-17 age group do have very high rates of rape and sexual assault. This age group was excluded to avoid issues with obtaining parental consent, as this would add cost to the survey and reduce the study's ability to achieve a high response rate. A separate study of sexual violence among juveniles should be initiated to work through not only the consent issues, but also development of more specialized instrumentation.

The address-based sample (ABS) was drawn from the U.S. Postal Service (USPS) address file. To efficiently screen for households, each sampled unit was mailed a survey asking for a listing of the age and sex of the members of the household. Returns from the survey were used to screen out households that did not have a female age 18-49 living at that address. All households that had at least one eligible person in the household or did not return the mail survey were followed up by a field interviewer. See Chapter 4 for more information on the execution of the mail survey.

Telephone Survey

A random digit dial (RDD) telephone survey was administered in the five CBSAs. This sample targeted women age 18 or older. The sample was not restricted to 18- to 49-year-olds because the cost of screening for a restricted age group by telephone was not an efficient way to meet the goals of the study. Also, by including all adults in the telephone sample, the analysis is able to generate estimates for all adults for at least one of the surveys. The RDD survey was a dual frame design, sampling from the universes of both landline and mobile telephone frames. To maximize the chances of contacting younger women, 80 percent of the sample was drawn from the mobile phone frame. In 2015 between 60 and 70 percent of the population age 18-45 did not have access to a landline, depending on the particular age group (Blumberg & Luke, 2016). Many of those who did have access to a landline did not use it as a regular mode of communication. For example, for the 2010 NISVS conducted 5 years earlier, only 6.9 percent of the landline surveys were completed with respondents age 18-29. This compared to 37 percent of the cell phone interviews (Black et al., 2011: 102).

The RDD frame drew the sample based on telephone numbers for the five CBSAs. The location in one of the CBSAs is primarily based on the area code of the telephone number. There are individuals who retain their cell phone number after moving. This means that the RDD sampling frame includes some individuals who did not currently live in one of the targeted areas (outmovers). There were also individuals who were not included in the RDD frame who lived in one of the areas but had an ineligible telephone number (inmovers). For cost reasons, the survey procedures completed interviews with the outmovers, and their data are included in the estimates described in this report. Outmovers were kept to represent inmovers who were missed. The assumption is that outmovers are similar to inmovers with respect to victimization. ZIP code information was collected during the interview, which made it possible to determine whether the respondent was an outmover. Using this information, rates of victimization were estimated for outmovers. Furthermore, rates for inmovers were estimated from the in-person sample by using the area codes from anyone who owns a cell phone. This was also collected during the interview. Chapter 10 analyzes these data and provides the proportion of inmovers and outmovers, as well as their respective victimization rates.

2.1.2 Volunteer Sample

Volunteers (VO) were recruited through ads on Craigslist in the five CBSAs aiming to recruit women 18 to 39 years old for a study on health and safety. For the RSA Pilot Test, the

Craigslist ads included a link to a brief web survey that was available in English and Spanish languages. Details on the process used to recruit this group are provided in Chapter 4.

2.1.3 Service Provider Sample

A fourth group of women recruited for the study were clients at rape crisis centers in the five CBSAs. As mentioned in Chapter 1, this group was recruited to collect information on how victims of rape and sexual assault would react to the survey procedures and the questionnaire. More information is provided in Chapter 4 on the methods used to recruit this group. Appendix K provides the results of the interviews conducted with this group.

2.2 Overview of survey procedures

The survey procedures were designed to maximize the coverage and response rate of the two surveys and address the human subject issues discussed in Chapter 1. The procedures differed slightly for each sample type.

2.2.1 In-Person Survey: General Population Sample

All initial contact for the survey was made by introducing the survey as one on “health and safety.” For the in-person general population sample, the initial contact with the household was the mail survey, which asked for an enumeration of the adults living in the household (see Chapter 4). Approximately 3,000 in-person interviews were completed.

All households eligible for an in-person visit were sent an advance letter describing the study, which was named the National Study on Health and Safety (NSHS). A field representative (FR) subsequently visited the household to select an eligible respondent. Once selected, the FR asked that the interview occur in a private setting, where no one could overhear or interrupt the conversation. The consent, which provided the purpose and content of the survey, was self-administered using the ACASI. Chapter 4 provides more details on recruitment, training, and monitoring of the FRs. Chapter 14 provides information on the extent to which FRs were able to establish and maintain a private setting during the interview.

Eligible respondents were offered \$20 to complete the interview. If an eligible respondent refused to do the interview, for whatever reason, no refusal conversion was attempted. This was out of concern with pressuring someone to participate who was reluctant to talk about her experiences.

2.2.2 Telephone Survey: General Population Sample

An RDD sample was used to draw a general population sample for the telephone survey. When initially calling into the household, the interviewer identified the survey as one on health and safety (see Appendix B for the telephone interview). Once a respondent was selected, she was told the purpose of the study and asked to go to an area where she would not be overheard before beginning the CATI. Just before the victimization screener was administered, the respondent was reminded of the importance for confidentiality and that she should feel free to hang up, if she felt in danger. All questions on the telephone survey were designed with either a yes/no response or by asking for a response to a number signifying the particular response category (with one exception). This was to make it difficult for anyone who did overhear the interview to understand what was being discussed.¹⁶ Eligible respondents were offered \$20 to participate in the interview. No refusal conversion was attempted for the same reasons noted above for the ACASI survey. Approximately 5,000 interviews were completed.

2.2.3 In-Person Survey: Volunteer and Service Provider Samples

For respondents in the volunteer and service provider samples who were assigned to the ACASI, FRs were instructed to call the potential respondent to set up a convenient time and location to meet. The preference was always to try to conduct the interview in the respondent's home, but if the home was not an appropriate location for any reason, the interviewer worked with the respondent to find a private and quiet alternative space. Approximately 1,000 volunteer interviews were completed with the volunteer sample and 17 with the service provider sample.

As with the general population sample, volunteer respondents were told during the recruitment process that the interview was about "health and safety." Respondents learned that the

¹⁶The one exception was at the end of the DIF where the respondent was asked to provide a description of the incident in her own words. Prior to getting a response to this question, the respondent was asked to confirm that she was in a private place where she could speak freely without being overheard by anyone.

interview was specifically about unwanted sexual experiences when they read the consent form. At this point, FRs were trained to answer questions about why the consent form focuses on sexual experiences when the study was previously described to them using different language. Appendix B provides the text used during the consent process.

In contrast, service provider respondents were told at the time of recruitment that the interview was about unwanted sexual experiences. To minimize the potential for respondent distress during the interview, potential respondents needed to be able to assess in advance their ability to handle an interview about these experiences. If the respondent indicated willingness to proceed with the interview, the interviewer scheduled a time and place to meet.

These respondents were offered \$30 to complete the interview, and the service provider respondents were offered an additional \$10 as reimbursement for any travel expenses. Interviewers made no attempts to convert refusals in either the volunteer or service provider samples.

2.2.4 Telephone Survey: Volunteer and Service Provider Samples

Volunteer and service provider respondents assigned to the telephone mode were called directly by the interviewer to introduce the study and request participation. The survey questions and procedures were identical to the general population telephone procedures. Approximately 1,000 volunteer interviews were completed with the volunteer sample and 41 with the service provider sample.

2.2.5 Re-Interview

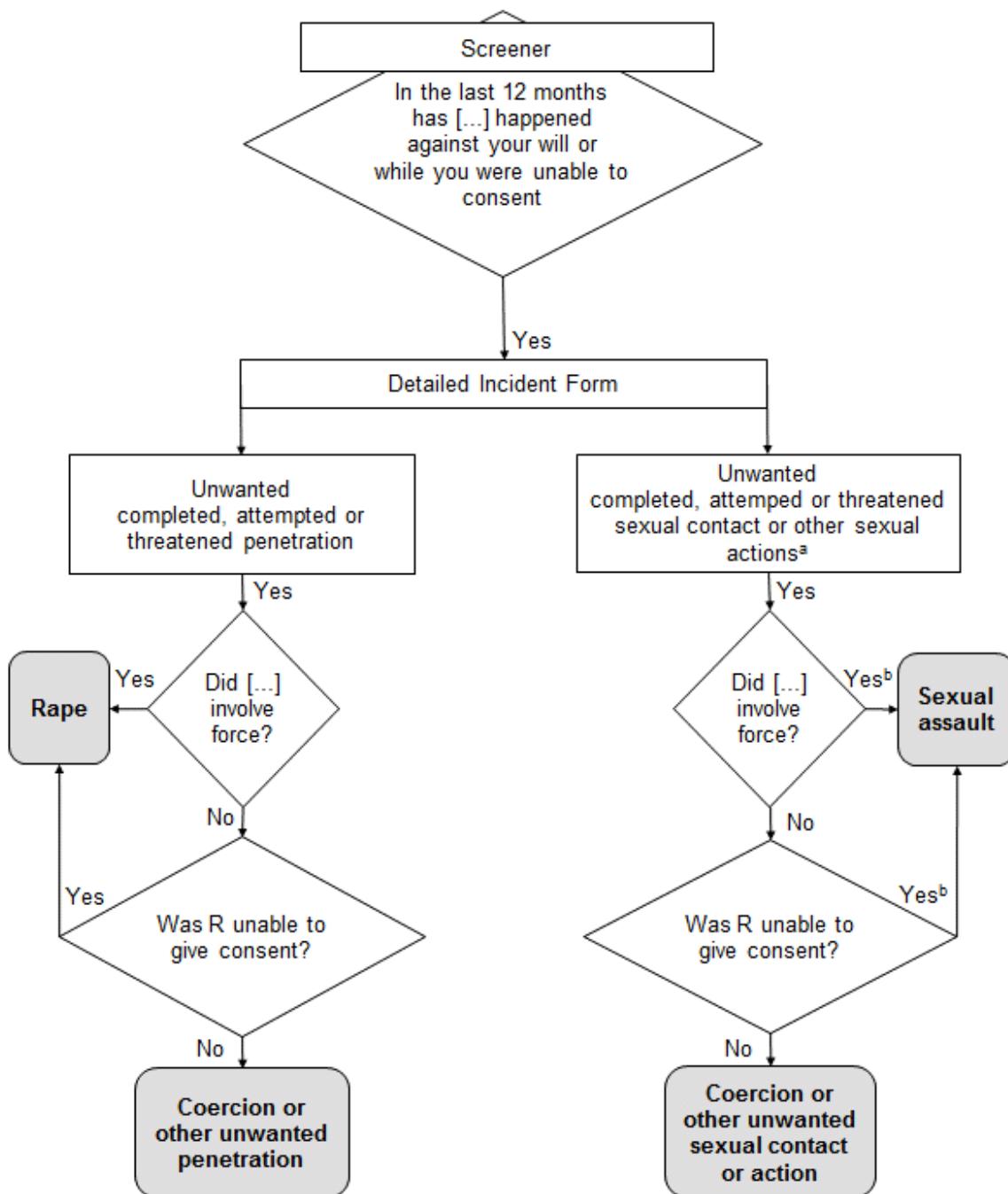
A subsample of the general population and VO respondents were re-interviewed about 2 to 3 weeks after the survey was completed. In total, approximately 1,000 re-interviews were completed. These were used as one method for evaluating the quality of the information collected for each mode of interview. The re-interview sampled respondents who reported a victimization at the first interview. The content of the re-interview was identical to the first interview. The goal was to estimate the reliability, and validity in some instances, of the estimates. Chapter 12 provides the details on the how the sample was drawn and the results.

2.3 Surveys to Measure Rape and Sexual Assault

Both the ACASI and CATI surveys used a two-stage design consistent with the current NCVS. The first stage included a series of BSQs asking about different types of unwanted sexual contact. The second stage collected details about the incident, including elements related to the tactics and behaviors. The intent was to count the number of unique incidents and to classify the incidents based on this information. The use of a two-stage design for the ACASI and CATI surveys allowed comparisons of estimates from the first and second stages for both modes.

The in-person and telephone instruments were designed to be identical in content. The content of the first and second stages were based on a review of the definitions of rape and sexual assault used in practice (e.g., by states and the federal government) and by other surveys. The BSQs used on the victimization screener were selected once a definition of rape and sexual assault was finalized. The logic of these definitions is illustrated in Figure 2-1 and discussed in the next section.

Figure 2-1. Criteria used to define rape and sexual assault



^aIncludes threatened sexual contact, exposure or video.

^bIf the offender stopped immediately, the incident was not classified as sexual assault.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

2.3.1 Defining and Screening Incidents as Rape and Sexual Assault

The RSA Pilot Test asked about four different types of sexual misconduct, including rape, sexual assault, coercive contact and unwanted contact. Papers commissioned by the National Academy of Sciences (NAS) panel provided guidance on the definitions for rape and sexual assault through a series of presentations (Tracy et al., 2013). A rape was defined as one of four different types of completed, attempted, or threatened penetration (vaginal, anal, oral, or digital) that occurred without the respondent's consent because of physical force or because the respondent was unable to consent. Sexual assault was defined as completed, attempted, or threatened non-penetrative sexual contact (e.g., kissing, groping, touching) or non-contact (e.g., exposure, exploitation by photos or video) by physical force or inability to consent. Coercive acts included sexual contacts where the offender threatened non-physical punishment (e.g., threats of job loss) or offered rewards (e.g., financial support, better grades). Unwanted sexual contact was defined as any other type of sexual contact the respondent said she did not want to have.

As shown in Figure 2-1, the essential elements of the definition of rape and sexual assault are the behavior and the tactic. The behaviors that make up rape are penetration of the vagina or anus with any body part or object, or oral penetration by a sex organ of another person. One tactic includes the use of physical force or threats of physical force. A second tactic is whether the person had the ability to consent. An incident may not involve physical force, but if the victim is unable to consent it was considered a rape. For a self-report survey of adults, perhaps the most important component for the ability to consent is impairment due to the use of alcohol or drugs. For example, all states now include statutes that cover situations when the victim is passed out (Tracy et al., 2013). Many states also include provisions where the victim is so impaired by alcohol or drugs that they are conscious, but unable to give consent.

The definition of sexual assault was based on the BJS definition:

“A wide range of victimizations, separate from rape or attempted rape. These crimes include attacks or attempted attacks generally involving unwanted sexual contact between victim and offender. Sexual assault may not involve force and include such things as grabbing or fondling. Sexual assault also includes verbal threats.”

In addition to including sexual contacts, the RSA Pilot Test incorporated other forms of noncontact sexual offenses, including exposure of sexual body parts and being forced to take photographs or videos.

To develop the screening items, the project reviewed a number of different surveys that have measured rape and sexual assault. The goal was to enumerate the targeted behaviors and tactics related to the definition of rape and sexual assault. The individual questions were selected to cover behaviors and tactics that were distinct and could not easily be combined into a single question. For example, separate questions were required to ask about incidents involving physical force and inability to consent. Similarly, separate questions were used for completed acts and attempted acts. It was also desirable to ask about related behaviors, such as those involving coercion or other unwanted contacts, to cue the respondent's memory for incidents that may qualify as a rape or sexual assault. For example, when asked about a sexual assault, a respondent may not immediately think of an eligible event, but when asked about a coercive contact her memory may be triggered to recall a sexual assault.

The final questions used on the screener (Table 2-1) were the result of a series of cognitive interviews to assess whether respondents understood the key concepts (see section 2.5.1 and Appendix L). The first four questions were based on the National Violence Against Women Survey (NVAWS; Tjaeden & Thoennes, 2000) and were intended to cover completed penetration of different types using physical force or threats of force. Four questions were used to target each type of penetration to make sure respondents understood the range of behaviors that were being targeted. Question 5 covers all types of penetration that occurred when the victim was unable to consent because of being unconscious and/or because of alcohol or drugs. This was the second tactic included in the study's definition of rape and was qualitatively different from the use of physical force. Question 6 was intended to cover coercion or non-physical threats. Question 7 covered instances of other types of nonconsensual situations, such as threats of harm to another person. Finally, question 8 was intended to cover attempted, but not completed, instances of penetration. Questions 9-14 cover other types of sexual contact including sexual touching, groping, and other attempted and completed types of sexual abuse.

The sources listed in the last column of table 2-1 provided the starting point for many of these items. These were used to guide both the behaviors and tactics. One modification to many of the questions was to express the nonconsensual nature of the act by the phrase "...against your will..." This was used for all items except when referring to being unable to consent because of being passed out or because of alcohol/drugs (question 5) and coercive acts (question 6). This phrase was used in selected items for the NVAWS. The NISVS uses the phrase "...when you didn't want it to happen..." which is somewhat broader. After cognitive testing, it was decided that the "against your will" conveyed the meaning of nonconsent in the most direct and specific manner (Appendix L).

Table 2-1. Questions used to screen for rape and sexual assault on the ACASI instrument

Q#	Wording	Source
1	Since [ANCHORDATE], has a male used force or threats of force to make you have vaginal sex against your will? By vaginal sex, it means putting his penis in your vagina.	NVAWS*
2	Since [ANCHORDATE], has anyone, male or female, used force or threats of force to make you have oral sex against your will? By oral sex, it means that someone penetrated your vagina or anus with their mouth or tongue, or you were forced to use your mouth or tongue on someone else's genitals or anus.	NVAWS*
3	Since [ANCHORDATE], has a male used force or threats of force to make you have anal sex against your will? By anal sex, it means that a man or boy put his penis in your anus.	NVAWS*
4	Since [ANCHORDATE], has a male or female used force or threats of force to put fingers or a foreign object in your vagina or anus against your will?	NVAWS*
5	(Other than the incidents you have already mentioned, since/Since) [ANCHORDATE] has anyone made you have any type of sex when you were unable to consent because you were too drunk, high or passed out?	NISVS*
6	(Other than the incidents you have already mentioned, since/Since) [ANCHORDATE], has anyone made you have any type of sex by threatening to cause problems for you, such as at your job or school, at home, in your relationships or in any other way?	NISVS*
7	(Other than the incidents you have already mentioned, since/Since) [ANCHORDATE], have you been in any other situations where someone made you have any type of sex against your will?	NEW
8	Thinking about all the different types of situations you have been asked about so far, since [ANCHORDATE], has anyone <u>tried, but did not succeed</u> at making you have any type of sex against your will?	NEW
9	(Other than the incidents you have already mentioned, since/Since) [ANCHORDATE] has anyone, male or female, kissed you in a sexual way against your will?	NISVS*
10	(Other than the incidents you have already mentioned, since/Since) [ANCHORDATE], has anyone, male or female, fondled, groped, grabbed, or touched you against your will?	NISVS*
11	Since [ANCHORDATE], has anyone <u>tried, but did not succeed</u> at kissing, fondling, groping, grabbing or touching you against your will?	NEW
12	(Other than the incidents you have already mentioned, since/Since) [ANCHORDATE], has anyone, male or female, made you watch against your will while they exposed their sexual body parts to you, flashed you, or masturbated in front of you?	NISVS*
13	(Other than the incidents you have already mentioned, since/Since) (ANCHORDATE), has anyone, male or female, made you show your sexual body parts to them against your will?	NISVS*
14	(Other than the incidents you have already mentioned, since/Since) [ANCHORDATE], has anyone, male or female, made you look at or participate in sexual photos or movies against your will?	NISVS*

NVAWS – National Violence Against Women Survey (Tjaden & Thoennes, 1998); NISVS – National Intimate Partner and Sexual Violence Survey (Black et al., 2011).

* Modified version of the question

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

A second difference from prior surveys was the order in which the screening items were presented. Some of the prior surveys have begun the screening by asking about non-penetrative acts first, in some cases psychological aggression (e.g., NISVS, NVAWS). In some surveys, rape is not asked about until after sexual touching and groping is asked (NISVS). This eases the respondent into

the most sensitive part of the interview. When interviewing victims of serious violence, this approach is also used to help respondents open up about what happened during the event. Unlike many of the prior surveys, the RSA Pilot Test started with the primary rape questions. This was done to reduce possible duplicate reporting across screener items. If sexual touching was asked first, victims of rape may wonder whether they should report the incident, especially if forced touching led up to the rape. On the other hand, victims of sexual assault that did not involve a rape may not be as likely to respond affirmatively to the rape question if it was asked first. This order then potentially reduces the number of times a respondent thinks she has to report on the rape during the screening process. Once reported at the first question, respondents should be less likely to report it in response to subsequent screening items. Since the NCVS is concerned with counting incidents, something other surveys have not prioritized as much in the past, this was considered an important advantage of the proposed order. Respondents were informed of the order at the beginning of the screening questions to prepare them for the first set of questions.

2.3.2 Detailed Incident Form

The topics covered on the DIF are shown in table 2-2. This also lists whether there were any items used in the crime classification. The algorithm used for the classification is provided in Chapter 7. The first section of the DIF confirms the incident that will be asked about and reminds the respondent that she can skip any questions if she does not want to answer (Sections A and B). Section C collected information on when and where the incident occurred. Section D collected a description of the type of unwanted behavior that occurred and the use of different tactics. One of the criticisms of using a DIF is that victims may see these as redundant with what was already reported in response to the BSQ. The respondent may even think the survey is questioning the veracity of what was reported on the BSQ. To reduce the likelihood of this, this section began by telling the respondent that

“...It may seem like you’ve already answered these questions, but we want to be sure we understand what happened to you during this incident.”

The first questions in this section ask about any unwanted behavior that occurred during the incident (see items D1 and D2 in Appendix B). These questions ask about each of the behaviors covered by the BSQs, including each type of penetration, kissing, sexual touching, exposure of body parts, and sexual photos or videos. The final question asked about any other behaviors that might have occurred during the incident. Each question asked whether the behavior was completed,

attempted, or threatened. To be classified as a rape or sexual assault, one of these behaviors had to be marked.

Table 2-2. Topics covered on the detailed incident form

Section	Topic	Used for classification?
A&B	Introduction and anchoring	
C	When and where did it happen	
D	Description of incident	
	Unwanted behavior	✓
	Coercive tactics	
	Physical force	✓
E	Physical injuries	
F	Offender characteristics	
G	Circumstances	
	Alcohol and drug use	✓
	Victim reaction	
	Offender reaction	✓
	Emotional distress	
H	Contact with service provider	
I	Follow-up with police	
	Consider it a crime?	
	Did police find out?	
	Anyone else told about the incident?	
	Contact any other organizations?	
Narrative		✓

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The next questions in section D asked about tactics that were used. Question D3 asked about non-physical tactics, including verbal pressure, financial threats, threats to cause problems in other life domains (job, school, relationships), or promising rewards of some type. Question D4 asked about different types of physical force or threats of physical force, including question about any other types of force not asked about. If nothing was checked in either D3 or D4, a question was asked if there was something else that made the event unwanted (D4f). If a respondent reported any type of physical force, they were classified as a victim of rape or sexual assault. The remaining questions in section D and E asked about the type of force used and physical injuries. Section F covered offender characteristics.

Section G collected information about the circumstances of the incident. Questions G4 through G15 cover the role of alcohol and/or drugs. If the incident involved the victim's use of a substance, she was asked whether she was passed out for all or part of the incident (G10). If she said "yes" to this item, the event was classified as a rape or sexual assault. In addition, if the respondent reported that she was unable to consent because of the alcohol or drugs (G12a), the incident was classified as a rape or sexual assault. The alcohol and drug items also collected information related to the respondent's condition, including her ability to remember the event (G11), whether she believed the perpetrator was trying to get her drunk to take advantage of her (G7), the effects on cognitive and physical functioning (G12a – G12d), and physical signs of being drunk (G13, G15). These variables are discussed in Chapter 8 to characterize the role alcohol or drugs played in the incident.

To better understand the context of the event and address the possibilities of false positives, section G includes items on the victim's reactions during the event, including how she communicated nonconsent. A follow-up is included on how the perpetrator reacted once the respondent said "no." The items on the victim's reactions were prefaced by a statement emphasizing that people react differently and that there were no right or wrong answers. The questions were being asked to get a better idea of how people react to these types of situations.

Sections H presents questions on some of the other consequences related to the incident, such as level of distress and contact with service providers. Section I asks about any follow-up with the police, whether the respondent thought the incident was a crime, the reasons why it was not reported to the police, and whether she had told any of her friends about the incident. The last part of section I measures whether the respondent contacted any organizations, offices, or agencies (e.g., victim service organizations) in the aftermath of the event.

To collect more information on the context of the incident, the DIF ended by asking for a description of the incident. This narrative was used in the analysis to check for accuracy of the responses to the DIF, as well as resolving differences between the screener response and the responses on the DIF. For respondents to the CATI, the interviewer first confirmed the respondent was in a safe and private location.

2.4 Structure of the Interviews

For general population respondents in both ACASI and CATI, the interview began with household screening questions to determine whether there was an eligible respondent living in the household. For ACASI households, this involved enumerating the sex and age of each adult resident age 18 or older; for CATI landline households, the Rizzo method was used (Rizzo et al., 2004) by assessing the number of male and female adults in the household; for CATI cell phone households, if the person answering the cell phone was male or under age 18, they were determined to be ineligible.

Apart from the event history calendar, which was presented via paper to in-person respondents to assist them in dating events and the tutorial for the ACASI system, the content of the interview was roughly equivalent across modes.

After obtaining consent to participate, the interview began with a short series of demographic and personal items. Next, the victimization screening questions (table 2-3) were administered to determine if the respondent had experienced any of 14 different types of unwanted sexual contact in their lifetime or in the past 12 months. The telephone protocol screened first for lifetime experiences, then determined whether a reported incident occurred most recently in the past 12 months. ACASI respondents were asked first if they had an experience in the past 12 months, and if not, were later asked if they have ever had such an experience in their lifetime (Appendix B). A discussion of results from the screening items appears in Chapter 9.

The different order of the lifetime and 12-month questions for ACASI and CATI modes reflects NCVS analytic priorities. The NCVS does not generate estimates of lifetime prevalence. Consequently, it was decided that the ACASI instrument should first ask about 12-month incidents. To get a lifetime measure and compare to the public health surveys, lifetime questions were then administered later in the screener if there were no reports to the 12-month question. As noted, the telephone survey starts with the lifetime period and then asks about 12-month incidents. This order was maintained for the telephone survey to keep the protocol as similar to methods that had been used on prior surveys.

Respondents reporting that one or more incidents occurred in the past 12 months continued into the DIF. The instrument was programmed to cap the number of DIFs to three incidents in the past 12 months, though respondents could be asked about up to five incidents if any of the DIFs

were incomplete. Priority was given to incidents involving rape, starting with the most recent incident. The content and order of the questions in the DIF was identical across modes. However, the CATI versions of the questions were structured to primarily ask for responses that were either yes/no or asking for the number of the response category. This was to protect the confidentiality of the respondent in case someone in the household was within earshot of the respondent.

Respondents providing data for more than one DIF were asked shortened versions of the DIF for the second and third incidents. A discussion of results from the DIF items appears in Chapter 8.

Following the DIF (or following the screener if no incident was reported in the past 12 months), the respondent was presented with two fictional vignettes that characterized different levels of coercion or alcohol use. For the alcohol vignette, respondents were randomly assigned to hear versions of the vignettes that varied the type of relationship, the drinking behavior, and the level of consent. The coercion vignette varied the type of relationship, the level of coercion, and whether or not “gentle force” was used. After being exposed to the vignette, respondents answered questions to determine the extent to which force or coercion was used in the incident. A discussion of results from the vignettes appears in Chapter 11.

After the vignettes, the respondent answered a short series of 10 debriefing items to assess their experience in completing the survey. These questions addressed distress and opinions about the survey, whether any portions of the questionnaire were difficult to understand, and in the re-interview, a short set of items to determine the utility of the resources provided at the end of the first interview. A discussion of results from the debriefing items appears in Chapter 14.

At the conclusion of the interview and after the field interviewer collected the laptop from the respondent, the interviewers checked the distress level of the respondent, and assuming it was safe to conclude the interview, offered local resources to the respondent, paid the respondent or collected her mailing address to mail the incentive, and, set up a time for the re-interview, if the respondent had been selected for a re-interview.

After departing the household or hanging up the phone, interviewers completed an observation questionnaire to record levels of cooperation, the level of privacy of the interview, and other observational factors. A discussion of results from the interviewer observation items appears in Chapter 14.

Table 2-3. Content of interview by mode of data collection

Survey content	ACASI		CATI
	CAPI	ACASI	
Household roster*	Yes		Yes**
Consent	Yes^		
Demographics	Yes		Yes
Event history calendar	Yes		No
ACASI tutorial		Yes	No
Victimization screener		Yes	Yes
Lifetime, then past 12 month		No	Yes
Past 12 month, then lifetime		Yes	No
Detailed incident form		Yes	Yes
Vignettes		Yes	Yes
Debriefing		Yes	Yes
Distress check	Yes		Yes
Provision of resources	Yes		Yes
Incentive payment	Yes		Yes
Re-interview request (if selected)	Yes		Yes

* Household roster was not administered for Volunteer or Service Provider samples.
**Separate household rosters were used for landline and cell phone samples.
^ Consent was administered via CAPI but the interviewer turned the screen to the respondent to read the content herself.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

2.5 Tests of Procedures and Surveys

Cognitive interviews and a feasibility test were conducted to help test and develop the procedures and surveys prior to launching the RSA Pilot Test. This section provides an overview of these activities. Appendices L and M provide more details on each.

2.5.1 Cognitive Testing

In February, March, and May 2013, three sets of cognitive interviews were conducted to test and refine the ACASI and CATI instruments that would eventually be administered in the Feasibility Study and Pilot Test. The first set of cognitive interviewing consisted of 23 women recruited from Craigslist and local colleges and universities in the Washington, DC, area. Westat recruiters screened the women to identify whether they had ever had any unwanted sexual experiences; 18 of the 23

women who ultimately participated in the interviews had experienced some unwanted sexual act in their lifetime, including 5 in the past 12 months. All respondents were asked to be re-interviewed 2 weeks later; 22 out of 23 women completed the re-interview. Additionally, 9 women were recruited from the DC Rape Crisis Center to participate in a cognitive interview.

Key findings from the first set of interviews showed that the DIF required greater refinements in how it asks about the types of assaults that were threatened, attempted, or completed; that greater detail was needed on the role of alcohol and drugs in the incident; and that de-duplication efforts in the screening questions were often confused by asking about lifetime estimates interleaved with past 12-month questions.

The second set of cognitive interviewing was conducted in May 2013. It consisted of 20 women who had experienced unwanted sexual experiences in their lifetime, with an emphasis on those having an experience in the past 12 months. Women were recruited from Craigslist and local colleges and universities in the Washington, DC, area. Results from these interviews revealed the need for a “catch-all” screening item to ask about other types of unwanted sex that may not have involved force, coercion, or alcohol/drug facilitation and suggested improvements in the measurement of frequent events such as unwanted kissing and groping; further refinements in differentiating between attempted, threatened, and completed acts; improved measures for levels of distress following the incident; and further modifications in measuring the role of alcohol and drugs in the incident.

The third set of interviews were conducted testing the Spanish version of the survey. These cognitive interviews were conducted in October 2013 after translation of the final instruments, allowing for a final set of adjustments to ensure the Spanish-speaking population understood the questions as intended.

Full reports from cognitive testing can be found in Appendix L.

2.5.2 Feasibility Study

Over a 4-week period in May 2014, Westat conducted a small Feasibility Study to test field operations and identify any outstanding issues prior to the Pilot Test. The Feasibility Study included a total of 92 CATI interviews and 97 ACASI interviews across the CBSAs in the Pilot Test—Los Angeles, Phoenix, Dallas, Miami, and New York City. Respondents were recruited from the general population, Craigslist asking for women ages 18 to 29, and a service provider sample recruited from local rape crisis centers. Both ACASI and CATI instruments were tested.

In addition, a total of 36 re-interviews were conducted, 18 in each mode. Respondents who indicated a past 12-month incident were prioritized for selection into the re-interview. Table 2-4 presents the results and completion rates for each sample type and mode.

Table 2-4. Completion rates for feasibility study household roster, main interview and re-interview by sample type and interview mode, 2014

Mode/sample type	Household roster				Main interview		
	Initial sample size	Eligible	Ineligible	Completion rate	Eligible	Completes	Completion rate
ACASI ^a	400	50	5	13.7 %	138	97	70.3 %
General population	400	50	5	13.7 %	50	38	76.0 %
Volunteers					80	52	65.0
Service provider					9	7	77.8
CATI ^b	1953	96	163	13.3 %	170	105	61.8 %
General population-Landline	650	59	6	10.0 %	59	32	54.2 %
General population-Cell Phone	1303	37	157	14.9	37	16	43.2
Volunteers					69	53	76.8
Service provider					5	4	80.0

Table 2-4 (continued)

Mode/sample type	Re-interview		
	Eligible	Completes	Completion rate
ACASI ^a	19	18	94.7 %
General population	7	6	85.7 %
Volunteers	12	12	100.0
Service provider			
CATI ^b	33	18	54.5 %
General population-Landline	5	1	20.0 %
General population-Cell Phone	2	0	0.0
Volunteers	26	17	65.4
Service provider			

Note: Estimates are based on unweighted data.

^aComputer-assisted telephone interview.

^bAudio computer-assisted self-interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 2-5 shows the approximate timings of the different sections of the Feasibility Study instrument as well as the overall timing. On average, the CATI instrument took 23.6 minutes to administer, whereas the CAPI/ACASI instrument took 21.7 minutes to administer. The major difference between the two instruments was that the DIF took longer when administered over the telephone.

Table 2-5. Approximate timings by section of the main interview by interview mode

Section	CATI (n=93)	CAPI/ACASI (n=92)
Consent	2.8	2.9
Demographics	4.3	3.4
Victimization screener	6.0	7.5
Detailed Incident Form (DIF)	42.7	16.9
Vignettes	4.0	2.6
Debriefing	1.2	1.5
Total No DIF	18.9	16.8
Total 1 DIF	40.9	34.6
Overall	23.6	21.7

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014.

Feedback from Feasibility Study respondents in the debriefing portion of the interview and from interviewers in a post-data collection debriefing indicated generally positive reactions to the survey. More than 8 in 10 respondents in both modes found the questions to be easy to answer, with the hardest questions being the vignette items. More than three-fourths of respondents indicated they were glad to have participated and would make the same choice to participate again. Interviewers enjoyed working on the survey and felt it gave respondents a “voice” on a topic that is not talked about openly.

As a result of the Feasibility Study, numerous changes to the survey and to study procedures were proposed. One major change made was to shorten the length of the DIF for those reporting a second and third incident. In addition, changes were made to item wording, translation, data collection procedures and materials, and interviewer hiring and training procedures. Each proposed change was prioritized, reviewed, and implemented if deemed high priority.

The full report from the feasibility study can be found in Appendix M.

3. Sampling and Weighting

This chapter describes the sample design and weighting of the two general population samples. Issues involving the sample design, the estimation methods, and the reliability of the final estimates were important during each phase of the study:

- During the planning phase, an essential question was whether the sample size and other resources would yield estimates of acceptable reliability to address the overall research questions. The sample design also needed to be sufficiently flexible to allow adjustments to the sample size during the field period while avoiding unacceptable impacts on reliability.
- During the implementation phase, the challenges of interviewing and securing an adequate response rate required modifications to the audio computer-assisted self interview (ACASI) sample design. The modifications were consistent with the goal of maintaining the validity of the probability sample.
- During the estimation and analysis phase, the original design and the subsequent modifications were reflected in the construction of survey weights. The weights compensate for unequal probabilities of selection for the sampled individuals, including the effect of the modifications to the sample design. The weights also align characteristics of the respondents with external information such as the age distribution of females in the core-based statistical areas (CBSAs). Equalized weights were created to treat each of the five CBSAs equally in the analysis rather than in proportion to their populations.

This chapter summarizes major elements of the sample designs for both the ACASI and computer assisted telephone interview (CATI) general population (GP) samples. The first section summarizes the planning phase, including the assumptions that were made to project nonresponse rates, to target sample sizes, and to estimate the reliability. The section also outlines the initial sample designs for the GP samples. Two sections follow describing the initial ACASI and CATI sample designs in more detail. The fifth section documents modifications to ACASI design during the implementation phase. The final two sections describe the calculation of survey weights used in the analysis and the assessment of reliability through the calculation of sampling variances consistent with the survey designs. The chapter is intended to provide a high-level overview of each of these topics.

3.1 Initial Sample Designs and Goals

The purpose of the study was reflected in the Bureau of Justice Statistics' (BJS) 2011 solicitation for methodological research, specifically research to support the National Crime Victimization Survey with respect to self-reported data on rape and sexual assault. The solicitation outlined the basic features of an "optimal design" and a "comparison design" and asked for an analysis of their relative performance. The solicitation set an upper limit of 10,000 completed interviews for each mode, and it directed that the final sample size be adequate to support the comparison between modes. "Up to 10,000 completed interviews will be conducted for the optimal design, with final sample size chosen to provide sufficient power and precision to observe change in key estimates" (U.S. Department of Justice, 2011).

The solicitation pointed to the potentially low number of reports of rape or sexual assault that would be observed with a sample size of 10,000 women, a projection drawing on both the National Crime Victimization Survey (NCVS) rates and findings from some other recent surveys. Historically, BJS has reported findings from the NCVS in the form of incidence rates, that is, the ratio of the estimated number of reported incidents in the previous year per 1,000 persons in the exposed population. Recently, BJS also began reporting prevalence rates, that is, the proportion of the exposed population that had been victimized during the previous 12 months by one or more incidents (in the BJS series *Criminal Victimization*, 2013 and 2014). Comparisons of NCVS to alternative surveys must account for whether the results pertain to incidence rates or prevalence rates and whether they apply to lifetime victimizations or victimizations during the previous year. The 1995-96 National Violence Against Women Survey (NVAWS) found a reported incidence of 8.7 rapes in the previous year per 1,000 women age 18 and over. By comparison, the 1996 NCVS obtained an incidence of only 2.3 rapes, attempted rapes, or sexual assaults per 1,000 women age 12 and over. Because the NVAWS found an average of 2.9 rapes per female victim, the estimated prevalence of rape was 0.3 percent, or about 3 victims per 1,000 women ($8.7/2.9 = 3$). Kilpatrick et al. (2007) projected a somewhat higher annual prevalence rate of 9 per 1,000 women. Using the sample size of 10,000, the solicitation remarked, "Extrapolating these rates to the target samples, the expected number of rape and sexual assault victims may range between 30 and 90, depending on the effectiveness of the screening strategies and types incidents covered in the surveys" (p. 12).

The solicitation invited consideration of alternative approaches to increase the observed number of victims interviewed, by seeding the sample with addresses of recent victims from police

reports or victim services. The solicitation did not specify the allocation between the seeded sample and the GP sample.

Rather than requiring an expensive national sample, the solicitation specified that the study be conducted in the following areas (core-based statistical areas, or CBSAs):

1. New York-Northern New Jersey-Long Island, NY-NJ-PA
2. Los Angeles-Long Beach-Santa Ana, CA
3. Miami-Fort Lauderdale-Pompano Beach, FL
4. Dallas-Fort Worth-Arlington, TX
5. Phoenix-Mesa-Scottsdale, AZ.

A list of the counties comprising each of the five CBSAs may be found in Appendix C1. The definitions are consistent with those used to publish the 2010 census.

The “optimal design” initially proposed included an in-person interview with an ACASI component for sensitive topics, while the comparison design was an RDD telephone survey with a proposed 25 percent drawn from a cell-phone frame.

In specifying the goals of the study, BJS directed that equal size samples for the RSA Pilot Test be selected in the five CBSAs, rather than sampling them in proportion to their population, as has been done for NCVS historically. In effect, separate samples were drawn for each CBSA, with each sample designed to be approximately the same size.

BJS’s decision to restrict the study to five large CBSAs, rather than require a national sample, simplified the sampling and data collection for the ACASI survey in several respects. For one, the primary sampling strategy was based on address-based sampling (ABS). ABS generally samples from address frames provided by commercial vendors who contract with the United States Postal Service (USPS) to repackage and market address information from the USPS Computerized Delivery Service Files (CDSF). Some vendors supplement the CDSF with information from other sources, including the vendor used for this study, Marketing Systems Group (MSG). The coverage of the vendors’ address files is generally good in urban areas but more limited in rural ones; the proportion of population in rural areas is quite small in the five RSA Pilot Test CBSAs. The Census Bureau uses a similar approach to sampling addresses for NCVS, but instead of commercial vendors, the Census

Bureau maintains its own address frame, including information from the decennial census and updates from the USPS and the American Community Survey (ACS).

One result of the design analysis was to clarify the target population for the two surveys. Although the solicitation limited the scope of the study to females, an analysis of expected costs led to an agreement to restrict the age range to 18-49 for the ACASI survey and to ages 18 and over for the CATI. The recommendation to set an upper limit of age 49 for ACASI was based primarily on the age-specific incidence rates for females in table 3-1, derived from NCVS public use files. The marked variation in rates by age is striking, and it factored heavily into design decisions. But given the obstacles of informed consent and access, the RSA Pilot Test design excluded minors younger than 18 from the target population, even though they show high reported rates.

Table 3-1. Incidence rate of rape per 1,000 females and standard errors by age group, National Crime Victimization Survey, 2005–2010

Age group	Rate	SE
12-17	3.7	0.5
18 or older	1.2	0.1
18-24	3.1	0.4
25-29	2.2	0.4
30-34	1.6	0.4
35-39	1.5	0.4
40-44	0.9	0.2
45-49	1.0	0.2
50-54	0.5	0.2
55-59	0.6	0.3
60-64	0.4 !	0.2
65-74	0.2	0.1
75-84	0.0 !	0.0
85 or older	0.3 !	0.3

Note: Estimates are based on weighted data.

!Interpret with caution. Coefficient of variation is greater than 50%.

Source: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2010

Differences in cost led to placing an upper age limit on the ACASI sample but not the CATI. In the ACASI sample, completion of many or most interviews required two steps: 1) obtaining a roster of the household to determine eligibility; and 2) scheduling a second interview with the sampled eligible respondent, if any, to administer the computer-based ACASI interview. The time required for the ACASI component contributed significantly to the overall cost. Investing a substantial portion of the ACASI sample to interview older adults with low incidence rates was inefficient in terms of the research goals. From an assumption that the relative cost of screening households was only half of the cost of completing the ACASI portion of the survey, the optimum efficiency in measuring rape and sexual assault occurred at or near the age groups 18-39. Expanding the age range to 18-49 only resulted in a 2 percent loss in efficiency, but also expanded the generalizability of the sampling frame.

In the CATI sample, previous experience supported an assumption that the cost of screening households would be roughly four times as high as the cost of completing the interview. The relative cost of screening was expected to be particularly high for the cell phone component, because cell phones were treated as a personal communication device not requiring creation of a household roster, so that once an eligible respondent was contacted, the main interview could proceed. Thus, completing the CATI cell-phone interview for an eligible female was expected to be of lower marginal cost than the second contact typically required for ACASI. The CATI cost per case was also projected to be lower than for the ACASI sample. For the CATI sample, the optimum efficiency occurred around 18-64, with roughly a 2 percent decline in efficiency by including all women age 18 and over.

The distinctively different optimum strategies for the two modes posed a choice. One option was to select a compromise age range, such as 18-54, so that the data collected for both modes would be analytically comparable. The compromise range would need to favor the ACASI optimum, given the higher cost per case for ACASI. The alternative selected was to use the individual optimum ranges, although selecting 18-49 for ACASI as a partial compromise. The CATI range was set to 18 and over. The decision recognized that CATI data on females age 50 and over would be useful for comparisons to other studies, even if this age range were only covered in one of the two modes. However, in this report, all mode comparisons between CATI and ACASI for the GP samples are based on ages 18-49.

A second recommendation was to increase the proportion of the CATI sample allocated to cell phones, from 25 percent to 40 percent. Although the cost per case of the cell phone sample was expected to be higher than the landline sample, the increase appeared to be cost effective in view of

the increasing proportion of younger adults in cell-phone-only households and the higher rates of victimization expected among young women.

The third recommendation was to expand the size and scope of the seeded sample. Initial investigations into the use of police records or victims identified by victim service agencies did not show promise for gathering as many as 500 volunteers for each mode. Instead of seeding the sample with known victims, the project recruited young women (age 18-29) within each of the CBSAs. This plan was based on the premise that very young adults experience high rates of sexual violence and would be an efficient way to collect data to evaluate the questionnaire, particularly the detailed incident form.

Given this revised plan, the size of the seeded portion of the sample was expanded from 500 to 1,000 people for each mode. The revision clarified the intention of analyzing the results as an experiment. Recruited participants would need to agree to cooperate in advance to either mode before their randomly assigned mode was known to them. The goal was to increase the numbers of reported incidents for methodological study of mode differences in the performance of the questionnaires.

Two of the three modifications potentially increased costs. The restriction to ages under 50 for the ACASI sample implied that more households would need to be screened initially. Based on data from the ACS, approximately 85 percent of households include one or more women age 18 or over, but only approximately 52 percent contain one or more ages 18-49. Second, the increase in proportion of the CATI sample allocated to cell phones would increase the costs per case. As a result, the goal for the ACASI GP sample was reduced to 7,500 overall. The logic for the reduction proceeded in two steps. In the first step, a reduction from 10,000 to 9,000 removed the projected seeded sample of 1,000 from both the ACASI and CATI GP samples. In the second, reductions from 9,000 to 7,500 interviews for the ACASI sample partially offset the increased costs of screening a larger number of households targeting females age 18-49. Calculations showed that about 25 percent more ACASI female victims ages 18-49 would be expected under this new approach than if 9,000 interviews for females ages 18 and over had been targeted. For the CATI, the reduction from 9,000 to 8,000 interviews accounted for the increase in the proportion of those being interviewed by cell.

With these refinements of the design, the next step was to estimate the expected reliability of the principal results. In addition to the sample size, the reliability would depend on the level of the reported victimizations in the survey. For example, if the survey found that women reported rape

and sexual assault at the same rates as found nationally in the current NCVS, then few cases would have been observed in the RSA Pilot Test. As a consequence, the relative standard error, expressed as a percent (the coefficient of variation [CV]), would have been quite large. A target reliability was proposed based on an expected number of reports of victimization.

As noted in the BJS solicitation, some studies of rape and sexual assault in the United States have found significantly higher 1-year victimization rates than the NCVS. Although NCVS results in table 3-1 were consulted for information about the age distribution of victims, other studies were used to estimate an overall level of victimization. Specifically, the estimates started from one of the rates cited in the solicitation, a finding of Tjaden and Thoennes (1997) of a prevalence rate for rape of 3 per 1,000 for females age 18 and over from the 1995-96 NVAWS. This rate was adjusted to a rate of 4.5 per 1,000 for females ages 18-49, taking the age distribution of table 3-1 into account. A rate for sexual assault excluding rape was then estimated as 7 times that amount, or 31.5 per 1,000, based on the approximate ratio of other forms of sexual assault to completed rape for females in the Crime Survey of England and Wales (CSEW). This survey recently implemented an ACASI design to measure sexual violence and was thought to provide the best available data. The combined projected rate for rape and sexual assault would be 36 per 1,000.

In addition to the size of the sample, the planned sample design is another factor affecting the expected reliability of the estimates. When a survey protocol requires interviewers to visit respondents in person, it is almost always advantageous to divide the frame into clusters or segments of units geographically close together (e.g., on the same block) to reduce travel costs. A sample of segments is drawn as one of the stages of sampling so that interviews may be conducted in a cluster of nearby households. The segments for RSA Pilot Test were basically census block groups, well-defined geographic entities typically including several hundred housing units.

The ACASI survey sample followed the typical practice of selecting a sample of households within each sampled segment. For the revised ACASI sample size of 7,500 completed interviews, the initial design was to sample housing units from 300 segments per CBSA, to yield five completed interviews per segment on average. Because of the restriction to ages 18-49, only about one out of two households was expected to include an eligible female, so it was necessary to sample enough households to account for the fraction of households with an eligible female, as well as for losses from nonresponse. The initial estimate was to sample 22 housing units per segment.

If a sampled household included more than one eligible respondent, one of them was selected at random. Females under age 30 were given twice the selection probability as other eligible

females. For example, a female under age 30 would have a two-thirds probability of selection if she lived in a household with one other eligible female age 30 or older. The survey weights include an adjustment for this subsampling within household, and the potential variation in weights was taken into account in projecting the expected reliability of the sample. The landline CATI sample faced a similar situation when more than one eligible respondent resided in the household sharing the same landline phone, and a similar selection of one respondent followed by an adjustment to the weights was employed, although simplified to give each eligible female an equal probability of selection.

A simple random sample of 7,500 women ages 18-49 would yield a CV of about 17 percent on the estimated prevalence of rape. To anticipate the effect of the clustered ACASI design, an interclass correlation of 0.05 was assumed to reflect the similarity of neighbors in a segment. For an average of five completed interviews per segment, the assumption predicts a design effect of about 1.2.¹⁷ The design effect due to clustering was further multiplied by an additional factor of 1.2, or a total design effect of 1.44, to reflect the variance impact of subsampling respondents in households where two or more females are eligible, and to allow for other possible sources of weight variation within each CBSA. As a result, the sample of 7,500 women from the ACASI design was projected to be equivalent to a simple random sample of 5,208 women ages 18-49.

Similar considerations were reflected in the projections for the CATI sample. Although geographic clustering is largely avoided with the CATI sample, anticipated variation in the weights from the dual landline and cell frames and other sources suggested using a design effect of 1.4 based on experience from other telephone surveys. In other words, the initial sample of 8,000 women age 18 or over was projected to yield 4,880 women age 18-49, equivalent to a simple random sample of about 3,429.

Table 3-2 summarizes the projected reliability for both the ACASI and CATI samples based on the revised sample sizes of 7,500 completed ACASI general population (GP) interviews and 8,000 completed CATI GP interviews. The projections indicated that estimates of the prevalence of rape over the preceding year would be relatively imprecise, with coefficients of variation (CVs) over 20 percent, but the prevalence of rape and sexual assault combined would be measured with CVs under 10 percent. As a result, the study would only reliably detect mode differences in reported rape by a factor of 2 or more, but would be more sensitive to differences in the overall rates for rape and sexual assault.

¹⁷The design effect represents the ratio of the variance under the clustered design compared to the variance under simple random sampling for a sample of the same size.

Table 3-2. Design assumptions for prevalence rates per 1,000 of rape and sexual assault over a 12-month period by interview mode for females ages 18-49 in the general population, 2014-2015

	Assumed value	
	ACASI ^a	CATI ^b
Rape and sexual assault ^c		
Prevalence rate per 1,000	36.0	36.0
Standard error	2.6	3.2
Coefficient of variation	7.2 %	8.8 %
Rape ^d		
Prevalence rate per 1,000	4.5	4.5
Standard error	0.9	1.1
Coefficient of variation	20.6 %	25.2 %
Sexual assault ^e		
Prevalence rate per 1,000	31.5	31.5
Standard error	2.4	3.0
Coefficient of variation	7.7 %	9.4 %

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^dIncludes penetrative sexual contact using force or while unable to consent.

^eIncludes non-penetrative sexual contact using force or while unable to consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Appendix C2 provides further details on the calculation of reliability and power analysis completed at this point in the planning. Appendix C3 compares the design assumptions in table 3-2 with the realized results.

3.2 Further Details on the ACASI Design

The relatively high rates of rape and sexual assault for women age 18-29 in NCVS suggested creating a separate stratum of blocks with high concentrations of either female students living in college dormitories or females age 18-29 living in households. The published counts from the 2010 census enabled the identification of these blocks, which were termed “stratum 1.” Specifically, stratum 1 included those blocks where either (1) the block included more than 20 female students in dormitories; or (2) the block had a population of 100 people or more, of whom at least 30 percent were females age 18-29 and where 10 percent or less of the population was in group quarters. The last restriction was added to avoid placing blocks into stratum 1 on the basis of a high group-

quarters population out of scope for the survey, such as in correctional facilities for women. Each block in stratum 1 was treated as a separate segment. The plan included sampling persons in stratum 1 at approximately twice the rate of stratum 2.

All remaining areas in the CBSA counties were classified as stratum 2. The geographic coding scheme used for the 2010 census was used to divide stratum 2 into segments. The census divides counties into tracts, tracts into block groups, and block groups into blocks. A typical RSA Pilot Test segment in stratum 2 was a single block group with stratum 1 blocks, if any, removed. In a small proportion of cases, block groups were combined to ensure a minimum size for the segment. In the final sample of 813 stratum 2 segments, only one of them included more than one tract, and only 11 others contained more than one block group.

A measure of size was first defined for each block to use in probability proportional to size (PPS) sampling of the segments. MSG, the sample vendor, provided block-level counts of addresses classified into different categories for all of the blocks in the five CBSAs. A housing unit equivalent estimate was formed by combining the counts of city-style addresses, vacants, units at addresses the Post Office classifies as drop points, and seasonal units, but excluding other counts such as P.O. boxes. The measure of size for each block was constructed by summing (1) the housing unit equivalent from the vendor and counts of (2) housing units and (3) women age 18-49 from the 2010 census. Rather than relying solely on the vendor's counts of housing unit equivalents, this compromise measure of size leveraged information from the 2010 census to reduce the risk of the vendor omitting a populated block. Incorporating the counts from the census into the measure of size was particularly useful in a few rural counties in the study where the vendor's address list fell well below the census counts. The census counts were also useful in stratum 1 blocks, where the vendor's address count was not meaningful because it typically did not reflect the dormitory population. The measure of size for a stratum 2 segment was the sum of the measures of size of its blocks.

In the initial sample, 300 segments were selected in each CBSA for a total of 1,500. Out of the 1,500, 25 were from stratum 1 and the remaining 1,475 from stratum 2. For stratum 1, a sample size of between three to seven segments was set within each CBSA by rounding to an integer 300 times stratum 1's proportionate share of the population measures of size within the CBSA. A PPS sample of stratum 1 blocks was then selected within each CBSA by systematically sampling from the frame sorted by whether the block included women in college dormitories.

The sampling of stratum 2 occurred in two steps. In the first step, all of the remaining blocks were formed into primary sampling units (PSUs). Most PSUs were single tracts, but small tracts were combined with adjacent tracts prior to sampling, and a few large tracts were split into two or more PSUs on the basis of block groups. PSUs were assigned measures of size in the same manner as segments, that is, by summing the measures of size of their blocks. Within each CBSA, the number of PSUs selected was set at 300 minus the number of stratum 1 blocks selected. The PSUs were assigned tract-level characteristics obtained from the ACS. Within each CBSA, a systematic sample of PSUs was drawn, where the PSUs were sorted by state, an indicator of central city, an indicator of 30 percent or more renters, an indicator of 20 percent or more poverty, county, and tract. In the second step, the PSU was divided into segments, and a single segment was sampled PPS from each of the sampled PSUs. The net effect was to produce a PPS sample of segments.

As a hedge against inability to complete the entire sample, the 1,475 segments of stratum 2 were divided into two waves, wave 1 and wave 2. Approximately five-eighths of the stratum 2 segments, 185 in each CBSA, were randomly assigned to wave 1. All stratum 1 blocks were included in wave 1. The addresses in wave 1 segments were sampled from a current address list timed to allow the sample to be drawn for release to the field at the intended start of interviewing. Wave 2 was deliberately delayed a few months, and the addresses were sampled from an updated address list in time for a possible release some months after interviewing of wave 1 began. The plan was to use results from wave 1 to estimate how much of wave 2 was needed, and a sample of segments could be selected from wave 2 for release.

Sampling rates for addresses within each sampled segment were set to give a self-weighting sample for stratum 2 within each CBSA. Preliminary calculations had suggested 22 addresses per stratum 2 segment, based on an 89 percent occupancy rate, a 70 percent screener response rate, a 52 percent eligibility rate (that is, the percentage of occupied households with one or more women age 18-49), and 70 percent ACASI interview rate for sampled eligible females, to yield an average of five completed ACASI interviews per segment. Prior to sampling, the target was increased to an average of about 26.5 housing units or approximately six ACASI interviews per segment. The increase reflected two considerations. First, the typical segments were now block groups containing hundreds of housing units rather than the smaller segments considered in earlier planning, possibly lowering the interclass correlation below 0.05. Second, the increase provided a hedge against a possibly lower response rate than initially projected. The within-segment sampling rate was doubled in stratum 1 blocks, including the sampling of any addresses of housing units also present.

Stratum 1 blocks with dormitory populations in the 2010 census were to be visited to identify college dormitories now present and to obtain a listing of dormitory rooms and, when possible, the usual number of occupants assigned to each. A PPS sample of rooms was to be selected and a single sample respondent was to be selected from each sampled room.

A small fraction of addresses in the CDSF of the USPS are drop points. Drop points represent two or more housing units without distinct unit designations. Although the CDSF has only a single address for the drop point, it also reports the number of units believed to be present. When the number of units at a drop point is small, the mail may be dropped on a table or through a common mail slot for the residents to sort. Addresses with a larger number of units may use staff to distribute the mail. The proportion of drop point varies geographically: Los Angeles has almost none, for example, but drop points are common in some parts of New York City. Drop points were accommodated in the RSA Pilot Test sample design by assigning measures of size to each address in the list within the segment. Individual housing units were assigned a measure of size 1, as were drop points with three or fewer units reported. Drop points with more than three units were assigned a measure of size equal to the number of reported units divided by 3. Interviewers were to list the actual number of units they found at drop points with measures of size greater than 1, that is, more than three reported units. Headquarters staff were to subsample the list at a rate 3 times the housing unit rate for the segment. For example, if the number of units at the drop point agreed with the drop point count on the frame, then three units would be selected. However, in the unusual instance that the measure of size for a drop point exceeded the number required to be sampled with certainty, the drop point was made a certainty hit and its units sampled at the housing unit rate for the segment. Probabilistically, units in drop points had the same chance of selection as other housing units in the segment, so no adjustments specifically for drop points were required later in the weighting.

In three rural counties on the fringes of the CBSAs, the counts of addresses from the vendor fell considerably below the 2010 census counts. Segments sampled from these counties were designated for field listing instead of relying on ABS. Because these counties included only a small fraction of their CBSA's total population, very few segments in these counties fell into sample. Separately, one stratum 1 block was designated for listing to resolve a puzzling conflict between the list supplied by the vendor and the census count.

Although the coverage of vendor-supplied lists in urban areas has been shown to be quite good, some ABS surveys incorporate coverage improvement operations to represent addresses

missing from the address lists. This option was considered for the RSA Pilot Test but dropped for reasons of cost and timing, as well as an expectation that the yield would be quite low.

3.3 Further Details on the CATI Design

The relatively rapid growth in the proportion of cell-phone-only households in the United States has necessitated the inclusion of a sample of cell phone users in telephone surveys of the general population. Persons in cell-phone-only households differ in several respects from those in households with landlines. Cell-phone-only households tend to be younger and poorer than households with landlines (Blumberg et al. 2014). Past research (Planty et al. 2013) suggests that these characteristics are correlated with the risk for victimization by rape or sexual assault. The RSA Pilot Test consequently allocated the majority of the telephone sample to the cell phone frame. Rather than the 40 percent referenced above, 80 percent was allocated to the cell phone frame. However, the RSA Pilot Test still needed to sample households with landline phones in order to represent the proportion of population with only landlines. Consequently, estimation methods for dual frame surveys were required to avoid overrepresenting households with both types of phones.

In a dual frame survey, households with cell phones can be sampled from the cell phone frame in two ways: (1) including just households with only cell phones (referred to as the screener method) or (2) including households with both a cell phone and landline as well as those with only a cell phone (referred to as the overlap method). In the first approach, sampled cell-phone respondents are asked if their household has a landline during an initial screening, and they are retained in sample only if they do not have one. The second method also requires asking whether their household has a landline, but the information is used in the estimation method to account for the overlap of the frames. Brick et al. (2011) showed that the overlap method has advantages over the screener method. Accordingly, the overlap method was used for sampling cell phone households for the RSA Pilot Test.

To sample households with landline phones, the method of random digit dialing (RDD) referred to as “list-assisted RDD sampling” was used. This method is designed to produce an unclustered sample with good coverage and efficiency (Tucker, Lepkowski, and Piekarski, 2002). List-assisted RDD sampling first specifies all possible 100-banks of telephone numbers that cover the particular geographic area of interest, where a 100-bank is defined to be the set of all possible telephone numbers with the same first eight digits (area code, exchange, and first two of the last four

digits). Next, each 100-bank is matched against published telephone directories to determine the number of listed telephone numbers in the 100-bank. Those 100-banks with at least one listed residential number are used to create the sampling frame for landline telephones.

In place of 100-banks for landlines, the cell phone frame was built from 1000-series banks of numbers dedicated to wireless service. The numbers of working 100-banks for landlines and 1000-series banks for each CBSA in this study are shown in table 3-3.

A major difficulty associated with creating a frame for cell phone sampling is that cell phone numbers are not readily associated with a degree of certainty to subnational areas, such as the five CBSAs in this study. Although 1000-series banks dedicated to wireless service are associated with geographic areas such as counties, the individual telephone numbers in them do not necessarily correspond to the county in which the owner of the number now resides, because a cell phone number generally corresponds to the location where the cell phone was purchased. For some owners, the relevant purchase may have been a cell phone some years ago, as owners replace the cell phone but keep their number. As a result, the cell phone sample is subject to greater degrees of overcoverage (inclusion of households not in the targeted geographic area) and undercoverage (exclusion of households from the targeted geographic area) than the landline sample.

To reduce the number of unproductive calls in the sample, the study used the Comprehensive Screening Service (CSS), offered by Market Systems Group (MSG). The service matches the landline telephone numbers to White and Yellow Pages of telephone directories to identify nonresidential business numbers in the sample. The numbers identified by this process were coded as ineligible and not released for telephone interviewing. CSS also applied an automated procedure in conjunction with manual calling to identify nonworking numbers. All sampled landline numbers, including those listed in the White Pages, were included in this test. Numbers found to be nonworking were coded as ineligible and not released for telephone interviewing.

The CSS also identified cell phone numbers that had been ported from landline exchanges where telephone customers are allowed to swap their landline phones for cell phones and keep the same number. The original landline numbers therefore would not fall in the cell phone exchanges used for sampling cell phones. For simplicity, ported telephone numbers were also removed from the sample, creating a small but additional source of undercoverage. (An alternative approach would have been to transfer the number to the cell phone sample for processing, but the resulting complexities were considered to outweigh the likely benefits.)

Table 3-3. Number of exchanges, working banks, and telephone numbers in the landline and cell phone sampling frames by core-based statistical area, 2014–2015

Random digit dialing sampling frame	Phoenix	Los Angeles	Miami	New York City	Dallas
Landline frame					
Number of working 100-banks ^a included in frame	35,881	116,571	63,577	201,851	58,798
Total number of telephone numbers in frame	3,588,100	11,657,100	6,357,700	20,185,100	5,879,800
Cell phone frame					
Number of 1000-series banks	5,527	18,637	8,205	29,672	8,952
Total number of telephone numbers in frame	5,527,000	18,637,000	8,205,000	29,672,000	8,952,000

^aIncludes working banks with at least one listed telephone number.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

In addition to CSS, address matching was performed using reverse directory services for landline sampled numbers. An advance letter was mailed with an associated address to encourage response at the screener stage.

It is not possible to prescreen cell phone numbers or obtain addresses for them using telephone directories, because such directories do not exist. Additionally, predictive dialing of cell phone numbers is prohibited by law (at least at the time of this study), so automated purging techniques for nonproductive or nonworking numbers cannot be used for cell phone samples. However, the sampling provider, Marketing Systems Group (MSG), was able to determine a cell phone number's activity status using their proprietary Cell-WINS procedure (Dutwin & Malarek, 2014). All numbers determined by this process to be "active" were dialed. Those classified as "inactive" or "unknown" were removed from the sample. Berzofsky et al. (2015) evaluated this practice for a survey in Ohio and found it was cost-effective and resulted in low undercoverage.

The result of the prescreening processes for both the landline and cell RDD samples is shown in table 3-4. For both landlines and cell phones, the number of records selected was deliberately larger to create a reserve sample. Instead of releasing the sample all at once at the start of interviewing, replicate samples were released as needed to meet the target interview goal.

Table 3-4. Number of records selected and dialed in the landline and cell phone sampling frames by core-based statistical area, 2014–2015

Sample type	Phoenix	Los Angeles	Miami	New York City	Dallas
Landline					
Number of records selected	14,393	12,821	15,000	8,634	15,000
Number of records dialed	3,660	3,895	5,689	3,519	4,549
Cell phone					
Number of records selected	30,400	34,600	33,300	29,700	27,400
Number of records dialed	19,746	23,562	24,112	18,927	19,151

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

3.4 Modifications to the ACASI Design During the Field Period

The wave 1 sample was released to the field in September 2014. By November, it became clear that the field costs would substantially exceed the budget. An inventory was made of stratum 2 segments in wave 1 that had not yet been assigned to interviewers or, if assigned, had not yet been

worked. These segments were pulled back to be held in reserve and made temporarily unavailable for assignment. In effect, the pulled back segments became a reserve sample, taking a role originally planned for wave 2. Because the segments where fieldwork had already begun did not represent a random sample of wave 1, a modification to the design was required to preserve the probability basis of the sample. Fortunately, the budget was sufficient to complete both the segments already begun and a sample of the reserved segments. A random sample was selected from the reserved segments and reassigned to the field. At that point, the reserve sample returned to the field could be used to represent the remaining reserved wave 1 segments. When in March 2015 it became clear that more segments could be assigned, a second sample of segments was selected from the reserved segments and also returned to the field. The final sample comprised 838 segments, all from wave 1, of which 813 were from stratum 2 and 25 from stratum 1. Table 3-5 accounts for the status of the wave 1 segments in the original sample.

In March, the second sample drawn from the reserved segments was stratified by the 10 field regions established for the RSA Pilot Test. The field regions were used to organize interviewer assignments. Four of the five CBSAs were each divided into two or three field regions. Estimates were made of the workload that could be completed given the remaining interviewer resources and incomplete assignments in each field region. Although not stratified by field region, the November sample had drawn segments from each field region as well. Table 3-6 presents the distribution of the reserve segments by field region. This information was incorporated into the ACASI weighting described in the next section.

Table 3-5. Assignment status of wave 1 segments, included and excluded from the final segment sample

Segment status	Phoenix	Los Angeles	Miami	New York City	Dallas
Total in wave 1	189	190	188	192	192
Total in final sample	169	163	156	178	172
Stratum 1	4	5	3	7	6
Stratum 2, started by November 2014	79	72	53	112	80
Stratum 2 reserve, returned to field in November 2014	40	43	61	28	40
Stratum 2 reserve, returned to field in March 2015	46	43	39	31	46
Stratum 2 reserve, excluded from interviewing	20	27	32	14	20

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 3-6. Status of the wave 1 segments initially withheld in November 2014, by field region

	Phoenix	Los Angeles	Miami	New York City			Dallas	
Field region	10	8	9	4	5	1	6	7
Segment status								
Sampled in November	40	31	12	24	37	3	6	19
Sampled in March	46	37	6	9	30	4	11	16
Dropped	20	11	16	20	12	0	1	13

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

3.5 ACASI Weighting

Two primary sets of weights were produced for the ACASI sample:

- The *population-based weights* reflected the probabilities of selection of respondents. They were adjusted to agree with external population controls from the ACS for each of the five CBSAs. The population-based weights indicate a larger number of females 18-49 in large CBSAs, such as New York City, than in smaller ones, such as Phoenix, in spite of their approximately equal sample sizes.
- The *equalized weights* were produced by adjusting the population-based weights in a CBSA by a single CBSA-level factor. When the equalized weights are used to analyze the survey, they produce identical estimates for the population of females ages 18-49 in each CBSA, thus placing each CBSA on equal footing. The equalized weights have been used in most of the analysis in this report.

The weighting process occurred in five stages:

1. calculation of segment-level base weights as the inverse of the probability of selection of the segment
2. calculation of household-level base weights for households and dorm-room level base weights for dormitories
3. calculation of person-level base weights
4. raking to ACS-based controls in each CBSA, producing the final population-based weights
5. equalization of the weights.

In the initial sample design targeting 7,500 completed interviews, 300 segments were selected in each CBSA. The number of sampled segments assigned to stratum 1 in each CBSA is shown in table 3-5, and the number in stratum 2 can be obtained by subtraction from 300. Thus, the sample size, n_{ps} , was fixed in advance for the five CBSAs, s , and the distinction, p , between stratum 1 and 2. The initial segment base weight, $ISEGW_{j(ps)}$ for segment j in sampling stratum ps is given by

$$ISEGW_{j(ps)} = \frac{\sum_{k \in (ps)} MOS_k}{n_{ps} MOS_j}$$

where MOS_j is the measure of size for segment j in sampling stratum (ps). For the segments in stratum 2, the wave 1 sample within each CBSA was then selected as an equal probability subsample with probability $wv_{(2s)} \doteq 5/8$.

The final base weight for a segment needs to reflect both the restriction of the final sample to wave 1 and the subsampling of the reserved segments within it. For stratum 2, the segment-level base weight is

$$SEGW_{j(ps)} = (frp_{j(ps)})^{-1} (wv_{(ps)})^{-1} ISEGW_{j(ps)}$$

where the final release probability, $frp_{j(2s)}$, was 1 for each segment that was already being worked before the November pull-back of the reserved segments in wave 1. For other segments in stratum 2, $frp_{j(2s)}$ is the fraction of segments in the reserve sample that were eventually released out of all the reserved segments in the field region. Table 3-6 provides the values for the calculation of $frp_{j(2s)}$. For example, in Phoenix, $frp_{j(2s)} = 0.8113 = (40 + 46)/(40 + 46 + 20)$. The values of $frp_{j(2s)}$ range from 0.5294 to 1.

In stratum 1, $frp_{j(1s)} = 1$ and $wv_{(1s)} = 1$, so that $SEGW_{j(1s)} = ISEGW_{j(1s)}$.

The household sample for wave 1 was drawn for all segments in July 2014, before the November pull-back of segments. In July, the sampling rate within segment was set to produce a self-weighting sample of housing units within the CBSA, based on $ISEGW_{(ps)j}$ rather than $SEGW_{j(ps)}$. In other words, the conditional probability p_{ij} of selecting housing unit i in sampled segment j was set to

$$p_{ij} = (w_{(ps)})^{-1} ISEGW_{j(ps)} \quad (3.1)$$

where for $p = 2$, $w_{(2s)}$ was the initially intended stratum 2 housing unit base weight within the CBSA for the initial sample. Consequently, the housing unit base weight, $HBWGT_{ij}$, is

$$HBWGT_{ij(ps)} = SEGW_{j(ps)} p_{ij}^{-1} \quad (3.2)$$

Differences between $ISEGW_{(ps)}$ and $SEGW_{j(ps)}$ result in modest variation within each CBSA among the housing unit base weights in stratum 2.

The value of the target weight $w_{(1s)}$ for stratum 1 was set to .5 $w_{(2s)}$ both for housing units and for rooms in dormitories. In this way, the sampling rate was doubled in stratum 1. Equation (3.1) also determined the within-segment sampling rate in stratum 1.

For households or dormitory rooms with more than one eligible respondent, a single respondent was randomly selected, with eligible respondents age 18-29 given twice the chance of selection as eligible respondents age 30 or over. If $p_{kij(ps)}$ denotes the conditional probability of selection of person k in sampled household i in sampled segment j , then the *person-level base weight* is

$$PBWGT_{kij(ps)} = HBWGT_{ij(ps)} p_{kij(ps)}^{-1}. \quad (3.3)$$

To reduce nonresponse and noncoverage errors, the person-level base weights were then calibrated to population totals estimated from the 2014 ACS public use files. The calibration method used for the RSA Pilot Test was raking, or iterative proportional fitting, an iterative procedure where the weights are proportionally adjusted to multiple sets (or dimensions) of marginal control totals in a sequential order. The iterations continue until the marginal sums of the raked weights agree with the corresponding marginal control totals for all set of margins, within a specified tolerance level.

In each of the five CBSAs, the demographic variables age, race/ethnicity, marital status, and education were used as dimensions for raking either alone or in combination with each other. The *final population-based weight*, $PWGT_{kij(ps)}$, can be expressed as

$$PWGT_{kij(ps)} = PBWGT_{kij(ps)} \prod_{d=1}^D RF_{dl} \quad (3.4)$$

where RF_{dl} is the raking factor for level l in dimension d where respondent k falls. The result of raking is to achieve almost complete agreement between the population control totals and the survey estimates of those totals using the resulting raked weights. The margins used for raking the weights are shown in Appendix Cces4 and C5 for the ACASI and CATI samples, respectively.

Because the five CBSAs vary substantially in size while the sample sizes are approximately equal, the weights from (3.4) also vary substantially. Equalized weights were produced by averaging the estimated population aged 18-49 over the five CBSAs and computing a single factor for each CBSA that adjusted the raked weights to agree with the average population for the 5 CBSAs. The equalized weights are given by

$$EWGT_{kij} = \frac{\sum PWGT_{k'i'j'}}{5 \sum_{(k'i'j') \in s} PWGT_{k'i'j'}} PWGT_{kij} \quad (3.5)$$

where the numerator is the sum of all final population-based weights and the denominator is 5 times the sum of the population-based weights in the CBSAs. Within a single CBSA, rates for the female population ages 18-49 using the final population-based weights will agree with rates using the equalized weights, because the leading ratio in (3.5) will cancel out. Equalization changes the overall estimates. However, the choice of the numerator in (3.5) is arbitrary, in the sense that the numerator cancels out in the calculation of any estimated equalized rate for the overall sample.

Tables 3-7 and 3-8 compare estimates based on the 2014 ACS with weighted estimates from the RSA Pilot Test for the five CBSAs combined and for the five individual CBSAs. Four columns are provided in table 3-7 for the overall comparison based on the combined estimates: the estimates are based on 1) the ACS, 2) the person-level base weights from (3.3), 3) the final population-based person weights from (3.4), and 4) equalized estimates based on (3.5). Comparison of the first two columns measures the extent to which an unadjusted probability-based weighting of the sample resembles the ACS, a recognized standard. Generally, the results are close. The base-weighted estimates do not indicate any systematically low response for the Hispanic or non-Hispanic black populations. The estimates exhibit relatively lower response for ages under 30, however, where the largest relative difference between the ACS and RSA Pilot Test is about 15 percent for 18- to 21-year-olds. Estimates by education and marital status are also close, with a slightly lower relative response by married women.

After the weights incorporate the raking adjustments to ACS controls, the estimates in the third column of table 3-7 closely agree with the first column. In particular, the adjustments bring the age distribution into close agreement. Minor differences appear; for example, the weighting process imputed missing demographic characteristics, but the imputations are not used in tabulation. Somewhat larger differences between the first and third columns appear for education, but these differences are present because the raking procedure adjusted education and marital status for ages 25 and over rather than for all ages.

The third and fourth columns of table 3-7 measure the impact of weight equalization on the distributions of demographic characteristics. Differences are relatively small, with minimal effect on the age distribution.

Table 3-7. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015

Characteristic	Overall			
	ACS estimate^a	Person-level base weights^b	Final population-based weights^c	Equalized weights^d
Race/Hispanic origin				
Hispanic	35.3 %	38.8 %	35.3 %	36.7 %
Black ^e	14.0	15.2	13.6	13.7
Other ^{e,f}	50.7	45.3	50.3	49.1
Age				
18 – 21	11.5 %	9.8 %	11.5 %	11.6 %
22 – 24	9.5	8.4	9.5	9.3
25 – 29	16.3	16.0	16.7	16.4
30 – 34	16.0	15.9	15.7	15.8
35 – 39	15.3	17.6	15.4	15.6
40 – 44	15.7	16.4	15.6	15.7
45 – 49	15.7	15.9	15.7	15.6
Educational attainment				
High school diploma, GED, or less	33.7 %	33.6 %	34.9 %	36.0 %
Some college	31.9	31.3	30.4	31.9
College degree or beyond	34.4	35.0	34.7	32.1
Marital status				
Married	42.1 %	44.0 %	42.4 %	43.3 %
Not married	58.0	55.8	57.4	56.6

Table 3-7. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015 (continued)

Characteristic	Phoenix			Los Angeles		
	ACS estimate^a	Person-level base weights^b	Equalized weights^d	ACS estimate^a	Person-level base weights^b	Equalized weights^d
Race/Hispanic origin						
Hispanic	33.8 %	38.6 %	33.6 %	47.7 %	49.6 %	47.7 %
Black ^e	5.5	5.2	5.5	6.4	7.0	6.4
Other ^{e,f}	60.7	56.0	60.5	45.9	43.3	45.7
Age						
18 – 21	12.0 %	8.4 %	12.1 %	12.2 %	10.1 %	12.2 %
22 – 24	9.7	9.1	9.6	9.9	9.5	9.9
25 – 29	15.9	19.7	15.9	16.6	15.8	17.1
30 – 34	16.3	16.7	16.3	15.7	17.2	15.2
35 – 39	15.9	18.4	15.9	15.0	16.2	15.2
40 – 44	15.6	16.3	15.6	15.4	17.2	15.2
45 – 49	14.6	11.5	14.6	15.2	14.1	15.2
Educational attainment						
High school diploma, GED, or less	35.9 %	36.4 %	36.8 %	35.6 %	35.2 %	37.8 %
Some college	37.2	32.9	35.6	33.4	30.9	29.5
College degree or beyond	26.9	30.7	27.6	31.0	33.9	32.7

Table 3-7. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015 (continued)

Characteristic	Phoenix			Los Angeles		
	ACS estimate ^a	Person-level base weights ^b	Equalized weights ^d	ACS estimate ^a	Person-level base weights ^b	Equalized weights ^d
Marital status						
Married	44.9 %	46.4 %	45.4 %	40.3 %	39.5 %	41.0 %
Not married	55.1	53.5	54.4	59.7	60.4	58.9

Note: Estimates are based on weighted data.

^aReflects population estimates from the 2014 American Community Survey (ACS) public use files.

^bReflects the probability of selection for a person within a sampled household within a sampled segment.

^cReflects the incorporation of raking based on respondent's age, race/ethnicity, marital status and education level.

^dReflects averaging the estimated population 18-49 over the five CBSAs and computing a single factor for each CBSA that adjusted the raked weights to agree with the average population for the 5 CBSAs.

^eExcludes persons of Hispanic or Latina origin.

^fOther race includes white, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, persons identifying as "other race" (CATI only), and persons identifying as two or more races.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

Table 3-8. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015

Characteristic	Miami			New York City		
	ACS estimate ^a	Person-level base weights ^b	Equalized weights ^c	ACS estimate ^a	Person-level base weights ^b	Equalized weights ^c
Race/Hispanic origin						
Hispanic	46.9 %	44.2 %	46.9 %	25.9 %	33.2 %	25.9 %
Black ^d	22.8	25.5	22.8	17.3	16.2	16.4
Other ^{d,e}	30.3	29.8	29.9	56.8	49.1	56.2
Age						
18 – 21	11.6 %	11.4 %	11.9 %	10.9 %	10.1 %	10.9 %
22 – 24	8.8	6.7	8.6	9.5	8.4	9.5
25 – 29	15.4	12.9	15.6	16.6	16.0	17.2
30 – 34	15.5	15.9	15.2	16.2	13.9	15.7
35 – 39	15.5	17.6	15.5	15.1	19.1	15.1
40 – 44	16.1	16.1	16.1	15.5	15.0	15.5
45 – 49	17.1	19.5	17.1	16.1	17.5	16.1
Educational attainment						
High school diploma, GED, or less	34.7 %	29.1 %	36.5 %	30.7 %	34.1 %	30.9 %
Some college	35.9	35.1	33.0	27.9	26.5	28.1
College degree or beyond	29.4	35.2	30.1	41.4	39.4	41.0
Marital status						
Married	38.8 %	43.5 %	40.1 %	41.3 %	44.1 %	41.1 %
Not married	61.2	56.5	59.9	58.8	55.4	58.6

Table 3-8. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, Miami, New York City, and Dallas core-based statistical areas, 2014–2015 (continued)

Characteristic	Dallas		
	ACS estimate^a	Person-level base weights^b	Equalized weights^c
Race/Hispanic origin			
Hispanic	29.2 %	27.9 %	29.2 %
Black ^d	17.2	22.4	17.2
Other ^{d,e}	53.6	49.6	53.4
Age			
18 – 21	11.1 %	8.4 %	11.1 %
22 – 24	9.0	8.2	9.0
25 – 29	16.0	16.5	16.1
30 – 34	16.5	17.3	16.4
35 – 39	16.1	16.4	16.1
40 – 44	16.3	18.4	16.3
45 – 49	15.0	14.9	15.0
Educational attainment			
High school diploma, GED, or less	36.2 %	33.0 %	37.9 %
Some college	33.4	35.9	33.2
College degree or beyond	30.5	31.1	29.0

Table 3-8. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for address-based sample respondents, before and after weighting adjustments, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015 (continued)

Characteristic	Dallas		
	ACS estimate^a	Person-level base weights^b	Equalized weights^c
Marital status			
Married	48.4 %	49.1 %	49.0 %
Not married	51.6	50.9	51.0

Note: Estimates are based on weighted data.

^aReflects population estimates from the 2014 American Community Survey (ACS) public use files.

^bReflects the probability of selection for a person within a sampled household within a sampled segment.

^cReflects averaging the estimated population 18-49 over the five CBSAs and computing a single factor for each CBSA that adjusted the raked weights to agree with the average population for the 5 CBSAs.

^dExcludes persons of Hispanic or Latina origin.

^eOther race includes white, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, persons identifying as "other race" (CATI only), and persons identifying as two or more races.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

The remaining sections of tables 3-7 and 3-8 compare estimates for the individual CBSAs. Only three columns for each CBSA are necessary, because rates computed with the equalized weights in (3.5) agree with rates using the population-based final weights in (3.4) at the CBSA level. The modest relative underrepresentation of ages under 30, particularly of those ages 18-21, appears in each of the individual CBSAs. Although the comparisons for other demographic characteristics exhibit somewhat more variation than for the overall sample, the general finding is that the sample estimates, even before the raking adjustments, manage to represent the demographic characteristics of the CBSA populations quite well.

The ACASI sample design was clustered, with only 838 segments containing the final sample of 3,053 completed interviews, or about 3.6 interviews per segment on average. To estimate sampling variances reflecting the clustering, weight variation, and effect of estimation, 160 replicate weights were produced for each observation, both for the population-weighted estimates and for the equalized weights. Most variances shown in this report for ACASI findings are based on the replicate weights for the equalized weighting. Additional detail is provided in Appendix C6 on the method to produce the replicate weights.

3.6 CATI Weighting

The weights produced for the CATI sample combine the data from the landline and cell phone samples into a single set of weights to estimate the two study populations of interest, both females age 18-49 and females age 18 or older, in the five CBSAs. The weighting process occurred in five stages:

1. calculation of household weights for the landline and cell phone samples separately
2. calculation of person-level weights for the landline and cell phone samples separately
3. compositing the landline and cell-phone person-level weights, followed by trimming them
4. raking the weights to agree with external estimates
5. equalization of the weights.

The CATI sample was initially stratified into the five CBSAs, s . Sampled phone numbers, j , in stratum s for phone sample type p were selected with probability n_{ps}/N_{ps} , where N_{ps} is the total number of phone numbers of type p in stratum s and n_{ps} was the number selected. Just as in the ACASI sample, the selection probabilities varied by stratum and type in order to target equal numbers of completed interviews in each CBSA. The selection probabilities also varied across phone type within a CBSA to achieve an 20/80 percent balance of landline to cellphone interviews. The resulting base weight for household j is simply

$$HBWGT_{j(ps)} = \frac{N_{ps}}{n_{ps}} \quad (3.6)$$

The base weight was then adjusted to remove numbers found to be nonresidential and to compensate for numbers with unknown residency status. At the end of data collection, most sampled phone numbers could be classified as either (1) residential (*res*) if they belonged to a person or household or (2) nonresidential (*nonres*) if they belonged to a business or had other nonresidential use. The remaining numbers were classified as unknown residency status (*unk*) if they yielded only an answering machine or a ring-no-answer result despite several call attempts.

Within each CBSA, additional poststrata, s' , were created. For the cell sample, the cases were classified according to whether they were called from a New York City or Dallas area code. For the landline sample, cases were classified according to whether the frame provided an address where an advance letter could be sent, as well as whether they were called from a New York City or Dallas area code. A total of 24 poststrata were defined. For each s' , the proportion of residential addresses among the known sample phone numbers was computed as

$$pres_{(ps')} = \frac{\sum_{j \in (ps') \cap res} HBWGT_{j(ps)}}{\sum_{j \in (ps') \cap (res \cup nonres)} HBWGT_{j(ps)}}$$

where the numerator represents a summation of the base weights over all known sampled residential numbers in the poststratum and the denominator is a summation over all sampled numbers with determined status. A preliminary household-level weight was then defined as

$$\begin{aligned} H1WGT_{j(ps)} &= pres_{(ps')} HBWGT_{j(ps)} \text{ if } j \in unk \\ &= HBWGT_{j(ps)} \text{ if } j \in res \end{aligned} \quad (3.7)$$

and 0 otherwise. A household-level weight completing the screener is then given by

$$H2WGT_{j(ps)} = \frac{\sum_{j' \in R_c} H1WGT_{j'(ps)} + \sum_{j' \in NR_c} H1WGT_{j'(ps)}}{\sum_{j' \in R_c} H1WGT_{j'(ps)}} H1WGT_{j(ps)} \quad (3.8)$$

where R_c is the set of screener respondents and NR_c the set of nonrespondents in screener nonresponse adjustment cell c with $j \in c$. Screener nonresponse adjustment cells, c , were defined within each CBSA by phone type, availability of an address for the phone number, and phone number treatment.

The CATI system automatically sampled a single eligible respondent from the $N_{j(ps)}$ eligible respondents at a landline number, even though the most frequent values of $N_{j(ps)}$ were 1 or 0. Therefore, for all households completing screeners, a person-level base weight for the sampled person k in household j is given by

$$P1WGT_{kj(ps)} = N_{j(ps)} H2WGT_{j(ps)} \quad (3.9)$$

An adjustment for extended interview nonresponse was computed separately for a set of adjustment cells, e . Respondents k to the extended interview in adjustment cell e received an adjusted weight given by

$$P2WGT_{kj(ps)} = \frac{\sum_{k' \in R_e} P1WGT_{k'j'(ps)} + \sum_{k' \in NR_e} P1WGT_{k'j'(ps)}}{\sum_{k' \in R_e} P1WGT_{k'j'(ps)}} P1WGT_{kj(ps)} \quad (3.10)$$

where R_e denotes the set of respondents and NR_e the set of nonrespondents in cell e . The cells e were formed within each CBSA by using telephone type p and age of the sampled respondent from the screener interview.

A composite weighting adjustment accounted for the overlapping component of the two samples, that is, women with both landline and cellular telephones. The person-level weights in (3.9) can be used to estimate the eligible population with cell phones or eligible females in households with landline phones, but they cannot be used directly to combine the two samples because of the potential overlap. The next step was to combine the two samples into one dataset and develop a

single set of weights (referred to as composite weights). The composite weight, $PCWGT_{kj(ps)}$, for person k in household j , is calculated as

$$\begin{aligned}
 PCWGT_{kj(ps)} &= \lambda P2WGT_{kj(ps)} \quad \text{for females from the landline sample with both types of} \\
 &\quad \text{phones} \\
 &= (1 - \lambda) P2WGT_{kj(ps)} \quad \text{for females from the cell sample with both types of} \\
 &\quad \text{phones} \\
 &= P2WGT_{kj(ps)} \quad \text{for females who were landline-only or cellphone-only}
 \end{aligned}$$

where λ is the compositing factor for respondents with both landline and cellular telephones from the landline sample. A compositing factor of $\lambda = 0.50$ was used.

Due to the accumulative effect of the weighting steps, a small number of cases received very large weights. Weight trimming was performed to reduce the variance impact of large weights. The weights, $PTWGT$ were trimmed to be no more than approximately four times the median weight in the respondents' CBSA. Weights were trimmed for 1.39 percent of the final sample.

Following a method similar to that used for the ACASI sample, the trimmed weights were calibrated to population totals estimated from the 2014 ACS public use files. To aid in comparisons of the CATI estimates to the ACASI estimates, the calibration separately affected the estimates for females 18-49 and for females 50 or older.

The same set of demographic variables was used as for the ACASI weighting, although implemented in a slightly modified manner. Again, age, race/ethnicity, marital status, and education were used as dimensions for raking. The raked weight, $PRWGT_{kj(ps)}$, can be expressed as

$$PRWGT_{kj(ps)} = PTWGT_{kj(ps)} \prod_{d=1}^D RF_{dl} \quad (3.11)$$

Table 3-9 provides the variables used in the raking.

A set of equalized weights was created for the CATI sample from the raked weights, similar to those for the ACASI sample in (3.5).

Tables 3-9 and 3-10 parallel tables 3-7 and 3-8. Again, comparison of the first two columns measures the extent to which an unadjusted probability-based weighting of the sample, in this case using the composite weights, resembles a recognized standard. The age distributions agree closely. The distributions by race and ethnicity and by marital status differ slightly, with married women underrepresented by a small amount. But the comparison indicates that the last of the four education categories is relatively overrepresented in the data set before raking. This category represents respondents with a 4-year college degree or higher. This discrepancy for education appears more substantial than observed for the ACASI sample.

The raking adjustments to ACS controls again produce high agreement between the first column and the third, as they had done for the ACASI sample. And when the third and fourth columns are compared, weight equalization again changes the distribution of the demographic characteristics by only small amounts.

The overrepresentation of the highest education category appears in each of the five CBSAs individually, although to somewhat varying amounts. Although the discrepancies are less pronounced, married women appear underrepresented in all five CBSAs as well as the overall sample.

For each completed case, 160 replicate weights were derived for variance estimation, both for the population-based weights and the equalized weights.

Table 3-9. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015

Characteristic	Overall			
	ACS estimate^a	Composite weights^b	Final population-based weights^c	Equalized weights^d
Race/Hispanic origin				
Hispanic	35.3 %	34.8 %	35.3 %	36.7 %
Black ^e	14.0	15.8	14.0	13.9
Other ^{e,f}	50.7	49.4	50.7	49.4
Age				
18 – 24	21.0 %	23.2 %	21.0 %	21.0 %
25 – 29	16.3	16.6	16.3	16.1
30 – 39	31.4	29.9	31.4	31.6
40 – 49	31.4	30.3	31.4	31.4
Educational attainment				
GED or less	12.5 %	11.2 %	13.2 %	13.6 %
High school diploma	21.2	17.9	20.6	21.1
Some college	31.9	30.3	30.6	32.7
College degree or beyond	34.4	40.6	35.6	32.6
Marital status				
Married	42.1 %	37.9 %	42.1 %	42.8 %
Not married	58.0	62.1	58.0	57.3

Table 3-9. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015 (continued)

Characteristic	Phoenix			Los Angeles		
	ACS estimate ^a	Composite weights ^b	Equalized weights ^d	ACS estimate ^a	Composite weights ^b	Equalized weights ^d
Race/Hispanic origin						
Hispanic	33.8 %	31.9 %	33.8 %	47.7 %	45.7 %	47.7 %
Black ^e	5.5	9.4	5.5	6.4	9.7	6.4
Other ^{e,f}	60.7	58.7	60.7	45.9	44.6	45.9
Age						
18 – 24	21.7 %	19.8 %	21.7 %	22.2 %	21.3 %	22.2 %
25 – 29	15.9	18.9	15.9	16.6	19.8	16.6
30 – 39	32.2	30.1	32.2	30.7	27.7	30.7
40 – 49	30.2	31.2	30.2	30.6	31.2	30.6
Educational attainment						
GED or less	13.0 %	13.5 %	14.9 %	15.5 %	13.6 %	15.5 %
High school diploma	22.9	19.8	19.9	22.0	18.7	22.0
Some college	37.2	34.9	37.9	30.5	26.7	30.5
College degree or beyond	26.9	31.8	27.2	32.0	41.0	32.0

Table 3-9. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015 (continued)

Characteristic	Phoenix			Los Angeles		
	ACS estimate ^a	Composite weights ^b	Equalized weights ^d	ACS estimate ^a	Composite weights ^b	Equalized weights ^d
Marital status						
Married	44.9 %	41.6 %	44.9 %	40.3 %	37.8 %	40.3 %
Not married	55.1	58.4	55.1	59.7	62.2	59.7

Note: Estimates are based on weighted data.

^aReflects population estimates from the 2014 American Community Survey (ACS) public use files.

^bReflects the probability of selection for a person within a sampled household, adjusted for the overlap for persons who have both cell phones and landlines.

^cReflects the incorporation of raking based on respondent's age, race/ethnicity, marital status and education level.

^dReflects averaging the estimated population 18-49 over the five CBSAs and computing a single factor for each CBSA that adjusted the raked weights to agree with the average population for the 5 CBSAs.

^eExcludes persons of Hispanic or Latina origin.

^fOther race includes white, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, persons identifying as "other race" (CATI only), and persons identifying as two or more races.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

Table 3-10. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015

Characteristic	Miami			New York City		
	ACS estimate ^a	Composite weights ^b	Equalized weights ^c	ACS estimate ^a	Composite weights ^b	Equalized weights ^c
Race/Hispanic origin						
Hispanic	46.9 %	44.5 %	46.9 %	25.9 %	28.4 %	25.9 %
Black ^d	22.8	20.2	22.8	17.3	19.1	17.3
Other ^{d,e}	30.3	35.3	30.3	56.8	52.6	56.8
Age						
18 – 24	20.5 %	18.5 %	20.5 %	20.4 %	26.3 %	20.4 %
25 – 29	15.4	18.3	15.4	16.6	14.1	16.6
30 – 39	31.0	28.6	31.0	31.3	31.7	31.3
40 – 49	33.2	34.7	33.2	31.7	27.9	31.7
Educational attainment						
GED or less	9.9 %	8.2 %	11.5 %	10.4 %	9.9 %	11.4 %
High school diploma	24.9	16.0	22.4	20.3	16.6	18.2
Some college	35.9	35.6	35.5	27.9	28.7	26.9
College degree or beyond	29.4	40.1	30.6	41.4	44.8	43.5
Marital status						
Married	38.8 %	38.1 %	38.8 %	41.3 %	35.7 %	41.3 %
Not married	61.2	61.9	61.2	58.8	64.3	58.8

Table 3-10. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015 (continued)

Characteristic	Dallas		
	ACS estimate ^a	Composite weights ^b	Equalized weights ^c
Race/Hispanic origin			
Hispanic	29.2 %	29.5 %	29.2 %
Black ^d	17.2	16.1	17.2
Other ^{d,e}	53.6	54.5	53.6
Age			
18 – 24	20.1 %	22.0 %	20.1 %
25 – 29	16.0	16.4	16.0
30 – 39	32.6	29.4	32.6
40 – 49	31.3	32.2	31.3
Educational attainment			
GED or less	13.9 %	12.5 %	14.6 %
High school diploma	22.3	20.8	23.2
Some college	33.4	34.8	32.6
College degree or beyond	30.5	31.9	29.7

Table 3-10. Comparison of American Community Survey estimates of demographic characteristics to weighted estimates for telephone sample respondents, before and after weighting adjustments, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015 (continued)

Characteristic	Dallas		
	ACS estimate ^a	Composite weights ^b	Equalized weights ^c
Marital status			
Married	48.4 %	42.7 %	48.4 %
Not married	51.6	57.3	51.6

Note: Estimates are based on weighted data.

^aReflects population estimates from the 2014 American Community Survey (ACS) public use files.

^bReflects the probability of selection for a person within a sampled household, adjusted for the overlap for persons who have both cell phones and landlines.

^cReflects averaging the estimated population 18-49 over the five CBSAs and computing a single factor for each CBSA that adjusted the raked weights to agree with the average population for the 5 CBSAs.

^dExcludes persons of Hispanic or Latina origin.

^eOther race includes white, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, persons identifying as “other race” (CATI only), and persons identifying as two or more races.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

4. Data Collection

This chapter describes the data collection methodologies of the two modes, in-person and telephone, for the three sample types: general population (GP), Craigslist volunteers (VO), and service provider volunteers (SP).¹⁸ A local network of field interviewers performed in-person data collection and telephone interviewers conducted the landline and cellphone data collection. The sample populations resided in five core-based statistical areas (CBSAs): Los Angeles, Phoenix, Dallas, Miami, and New York City. The start and end dates of data collection periods by mode and sample type are shown below in table 4-1.

Table 4-1. Data collection periods

Sample type	Mode	
	In-person (ACASI)	Telephone (CATI)
GP	10/27/2014 to 6/30/2015	5/26/2015 to 10/11/2015
VO	10/13/2014 to 6/30/2015	10/20/2014 to 2/01/2015
SP	10/13/2014 to 6/30/2015	10/20/2014 to 2/01/2015

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Field interviewers completed on average 456 in-person main interviews per month during a 9-month period. Telephone interviewers completed on average 643 interviews per month during 10 months of data collection activity as shown in table 4-2, below.

4.1 In-Person – General Population Sample

The in-person survey involved several different steps. The first was to complete a household roster via a mail survey. The second was to follow-up the non-responding households to the mail survey with an in-person visit to complete a household roster. The third step was to complete the extended interview with the selected respondents.

¹⁸Discussion of the results for the SP sample are provided in Appendix K.

4.1.1 Household Screening – Roster Mailing

The first stage of in-person data collection of the GP sample involved screening households to identify those with eligible residents (females ages 18 to 49). Each address in the GP sample was sent a series of mailings in August and September 2014. The initial mailing packet included a cover letter that explained the purpose of the study, \$2 in cash, a postage-paid return mailer, and a household roster. (See Appendix J, Household screening by mail roster.) For these mailings, the study was titled “The National Study on Health and Safety.” These materials did not specify the topic of the survey.

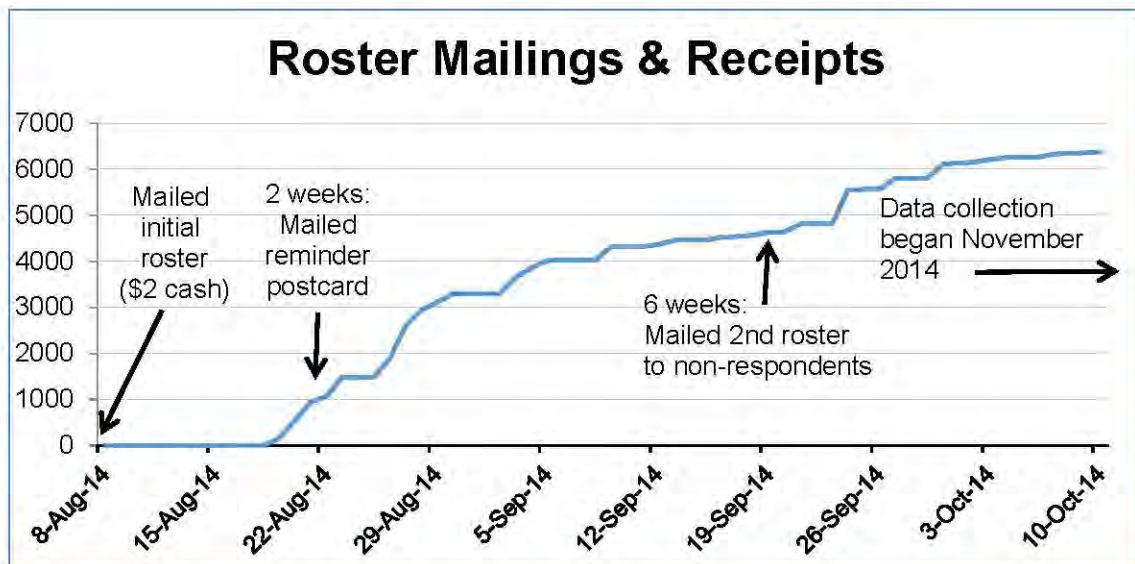
Table 4-2. Completed main interviews by month

	ACASI				CATI				
	GP	VO	SP	ACASI total	GP cell	GP landline	VO	SP	CATI total
Oct-14	-	127	-	127	-	-	76	1	77
Nov-14	519	418	4	941	-	-	495	6	501
Dec-14	414	169	3	586	-	-	465	1	466
Jan-15	347	83	3	433	-	-	100	1	101
Feb-15	321	41	4	366	-	-	-	-	-
Mar-15	428	25	6	459	-	-	-	-	-
Apr-15	389	31	17	437	-	-	-	-	-
May-15	351	104	4	459	55	12	-	-	67
Jun-15	284	14	-	298	846	86	-	-	932
Jul-15	-	-	-	-	1,005	44	-	6	1,055
Aug-15	-	-	-	-	1,387	580	-	2	1,969
Sep-15	-	-	-	-	872	210	-	-	1,082
Oct-15	-	-	-	-	133	50	-	-	183
Total	3,053	1,012	41	4,106	4,298	982	1,136	17	6,433

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The actual topic of the survey was shared when administering the informed consent to the selected respondent. All addresses received a reminder postcard 2 weeks after the initial mailing. Four weeks after the postcard mailing, all non-responding addresses received a second roster packet (without incentive). The elapsed times between the mailings and the returned rosters are shown in figure 4-1 below.

Figure 4-1. Roster mailings and receipts



Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Of the 24,607 rosters mailed, 6,489 households returned usable rosters, for a 26.4 percent return rate (table 4-3, below). Of these, 2,734 rosters indicated that an eligible respondent lived in the households and 3,755 indicated no eligible respondents in residence.

A follow-up letter was mailed to eligible households informing them that they were eligible for the full study and alerting them that a field interviewer would be contacting them. The addresses were excluded from further data collection activity when the USPS returned both mailings as "undeliverable," indicating no eligible respondents were in the household. The remaining addresses (i.e., addresses with eligible respondents, non-responding addresses, and addresses with one undeliverable from the USPS) were designated for field data collection. Households with eligible respondents received a letter informing them that an interviewer would stop by in the near future. A portion of these addresses were systematically de-selected and removed from the sample. This

exclusion is described in Chapter 3, section 3.4.¹⁹ The final sample yielded for the in-person collection was 18,126 addresses. These were sent to field interviewers so interviews could start the week of October 27, 2014.

Table 4-3. Outcome of ABS sample roster mailing by disposition

	ABS sample	Percent	Disposition
Rosters mailed	24,607		
Returned usable rosters	6,489	26.4	
Eligible respondent in household	2,734		Fielded
No eligible respondent in household	3,755		Ineligible household
Returned unusable rosters	261	1.1	
Incomplete	74		Fielded
Refusals	73		Fielded
Blank	114		Fielded
U.S. Post Office returns	1,868	7.6	
Both mailings non-deliverable	991		Out of scope (vacant or non-residential address)
Single mailing non-deliverable	877		Fielded
Non-response	15,989	65.0	Fielded
Final outcome			
Total finalized at roster mailing stage	4,746	19.3	
Total sent to field for in-person visit	19,861	80.7	
Total	24,607	100.0	

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

An experiment was conducted in the initial mailing to test the effectiveness of sending one \$2 bill vs. two \$1 bills. The results of this experiment showed that the rosters sent with a \$2 bill yielded a significantly better response rate (30.7% vs. 23.4%) over those with two \$1 bills ($p < .0001$). Details about the methodology and results of this experiment are provided in Appendix D.

4.1.2 Household Screening – In-Person Visits

The field data collection period for the GP ACASI respondents extended from late October 2014 through June 30, 2015. Field interviewers were responsible for locating the households and completing the in-person household roster. The field interviewer completed the household roster for

¹⁹In the beginning of December, 6,837 GP addresses in 318 segments were made inactive. In March of 2015, 4,409 of these cases were released back to the field.

all GP addresses, excluding those that had returned a mail roster showing an ineligible respondent. The in-person roster was administered by computer-assisted personal interview (CAPI). When one or more eligible respondents lived in the household, the CAPI program selected the respondent for the main interview. Field interviewers assigned a final disposition to households with no eligible respondents, that refused participation, or were otherwise unable to participate. Field interviewers made five in-person visit attempts to contact residents at each address to complete the household roster.²⁰

4.1.3 Respondent Interviews

Field interviewers administered the main interview survey to respondents via laptop computers using a combination of computer-assisted personal interviews (CAPI) and audio computer-assisted self-interviews (ACASI). Before starting the main interview, field interviewers first made sure the interview could be conducted in a place that was out of earshot from anyone else in the household. At that point, the sampled person was administered the informed consent using the ACASI and was informed about the survey and the confidential nature of the questions. The sampled person used the ACASI to indicate agreement to participate in the ACASI. Field interviewers conducted the interview where their conversation could not be overheard and the laptop screen could be seen only by the respondent during the ACASI portion of the survey. During the interview, the field interviewer addressed any questions or issues and monitored the respondent for signs of distress. (See Appendix D for the distress monitoring protocols) A total of 3,053 GP respondents completed the main interview. Of these, 362 (11.9%) completed the interview using the Spanish language instrument.

4.1.4 Respondent Re-Interviews

At the end of each main interview, the CAPI computer program selected a portion of the respondents to be re-interviewed. If the respondent agreed, the field interviewer scheduled an appointment approximately 2 weeks after the original interview. The re-interview was an exact repeat of the main interview and used the same anchor date and reference period as the initial interview.

²⁰During the final 6 weeks of data collection, addresses with as many as three visit attempts were closed out as nonrespondents in order to prioritize field activity for the remaining field period on addresses with 0 to 2 attempts.

4.1.5 Respondent Incentives

Respondents who completed the main interview or re-interview received a \$20 debit card as an incentive for their participation. The respondents were provided with a toll-free telephone number that would connect them with project staff, who provided assistance and answered questions regarding the debit card.

4.1.6 University Sample

A small set of segments encompassed residential housing units at 19 colleges and universities. Segment maps were developed using the address of the units and compared to university housing maps obtained online to identify all of the housing units within the segment map boundaries. Offices of the university presidents, chancellors, or provosts were contacted and received a follow up with a letter from BJS and brochure that explained the study. Recruitment efforts obtained cooperation at 12 of the 17 institutions that were in the sample. The cooperating institutions are shown in table 4-4.

Table 4-4. Participating colleges and universities

CBSA	Institution
AZ	Thunderbird School of Global Mgmt.
AZ	Arizona State University
AZ	Midwestern University
CA	University of Southern California
FL	Johnson and Wales University
FL	University of Miami
NY	Barnard College
NY	Manhattanville College
NY	New York University
TX	Texas Christian University
TX	Texas Woman's University
TX	Christ for the Nations

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Universities and college that declined participation cited timing conflicts with other studies taking place on campus and the need to avoid disruptions to the youth.

Using detailed maps and housing lists provided by university housing authorities, occupied rooms were identified. Housing directors helped to coordinate communications, establish contact protocols, and schedule field interviewer visits to the housing units. The university provided the

name of the residential unit, total housing capacity, unit numbers, room capacity, and layout of selected resident facilities. A roster collected information about the residents and a packet of information about the study was provided to the selected housing units. The packet of information, typically delivered by university housing staff, included an introductory letter, \$2 in cash, the roster (see Appendix D), and a study flyer indicating when interviewing personnel would be on campus to review the residents' eligibility for the study.

On most campuses, field interviewers were not permitted direct access to housing rooms due to campus security. Working with the office of the university housing director, field interviewers set up tables in common areas on campus to follow up with the residents who received the information packet and to solicit participation in the study. In a few occasions where the housing units comprised apartment-style rooms, the university allowed field interviewers to contact residents directly by knocking on doors of the selected rooms.

Once the interviewer established a resident's eligibility, an interviewing appointment was arranged. Most universities provided a private location on campus to conduct the main interview. A thank-you letter was mailed to the universities expressing appreciation for their support and assistance in making the study successful. The results of the university data collection are shown in table 4-5.

Table 4-5. University sample results

	Phoenix	Los Angeles	Miami	New York City	Dallas	Total
Institutions	3	1	2	3	3	12
Sampled living units	159	21	47	96	77	400
Unoccupied rooms/no eligible respondents	60	0	0	21	18	99
In-scope living units	99	21	47	75	59	301
Completed roster – ineligible	23	6	5	8	0	42
Completed roster – eligible	68	10	35	51	26	190
Non-response/refusals (eligibility unknown)	8	5	7	16	33	69
Roster completion rate	92%	76%	85%	79%	44%	77%
Completed main survey	63	9	20	40	23	155
Eligible non-response/refusals	5	1	15	11	3	35
Main survey completion rate	93%	90%	57%	78%	88%	82%
Main survey response rate	64%	43%	43%	53%	39%	51%

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

4.2 Telephone – General Population Sample

The telephone sample was a dual-frame random digit dialing (RDD) sample of landline and cellphone numbers. The interview was conducted using a computer-assisted telephone interview (CATI) mode to administer the survey. The design called for 4,000 completed interviews to be obtained from the cell phone sample, as well as 1,000 completed landline sample interviews. The RDD phase of data collection started on May 26, 2015, and was completed 20 weeks later, on October 11, 2015. The landline phone numbers that could be matched to mailing addresses were mailed an advance letter 1 week prior to the start of phone calling. (See Appendix D.)

The remainder of this section describes the steps taken to complete an interview. This includes completing an initial screening interview to determine eligibility, selecting a respondent, and completing the main interview.

4.2.1 Screening

There was a separate screening procedure for each of the two RDD sampling frames. For the landline telephones, an adult female was randomly selected among those living in the household. The initial call began by describing the call as a scientific study on health and safety sponsored by the U.S. Department of Justice. Household members were told that the study is seeking eligible respondents, noting that those selected would receive \$20 after completing the study. The Rizzo selection method (Rizzo et al., 2004) was used to select an adult female.

For the cell phone sample, the phone was considered an individual device, not shared by any other person. Consequently, there was no need to randomly select a respondent within the household. However, it was necessary to screen the individual for sex and age. When calling a cell phone, the telephone interviewer introduced the survey using a similar introduction as used with the landline. Before introducing the study, the interviewer first asked whether the person answering was driving or engaged in another activity that required their full attention. If not, the telephone interviewer asked for the sex (female) and age (at least 18 years old) of the person answering the phone. If the person was eligible, the interviewer administered the main interview.

Up to seven calls were made to a phone number with which no contact was made (ring no answer or answering machine results). The CATI scheduler algorithm ensured the calls were made

on different days of the week, and during different times of day, for each phone number. If no contact was made by the seventh call, the case was retired. For the first two and a half months of production, such cases were put back out for a second round of call attempts after a week's hold period – up to seven additional calls could be made, after which the case would be retired and considered non-response.

4.2.2 Respondent Interviews

At the extended interview level up to nine call attempts were made. If no contact was achieved, the case was retired as non-response. During the first two and half months of the project, these retired cases were put back out for additional calls after a short period of time.

When the telephone interviewer reached an eligible respondent, they requested permission to record the interview and walked the respondent through the informed consent. (See Appendix D.) The sensitive nature of the study was revealed only to the respondent during this stage as confidentiality and safety protections. If another member of the household completed the screening, that person knew only that the study concerned health and safety.

When the respondent consented to begin the survey, the telephone interviewer ensured that the respondent was in a private location and conducted the interview. Response options to questions used generic indicators such as numbers, letters, or yes/no responses so no one who overheard the respondent would know the nature of the questions. As a further protection, the telephone interviewer informed the respondent that she could terminate the interview at any time by simply saying goodbye and hanging up. At the start of the interview, the telephone interviewer also provided a study-specific, toll-free telephone number so the respondent could contact the study for any reason (e.g., to schedule a time to resume an interrupted interview). As described above for the in-person re-interviews (section 4.1.3), the CATI program selected a small portion of the respondents for an identical follow-up interview to be conducted 2 weeks later. A total of 5,187 general population respondents completed the main interview (965 from the landline sample and 4,222 from the cell phone sample). Of these, 650 (12.5%) completed the interview using the Spanish language instrument, 60 among the landline sample (6.2%), and 590 among the cell phone sample (14.0%).

4.2.3 Respondent Incentives

All RDD respondents were offered a \$20 incentive if they completed the interview. This offer was made for both the initial interview and the re-interview. The incentive was sent by check to the name and address provided by the respondent.

4.3 Volunteer Sample

The volunteer sample was recruited to supplement the general population sample. The goal was to recruit individuals who had a higher than average risk of victimization and were used to add information on how victims would respond to critical parts of the questionnaire.

4.3.1 Recruitment

In late summer 2014, volunteer sample (VO) recruitment began through Craigslist in the five CBSAs. By early January 2015, a sufficient number of volunteer (VO) recruits had been obtained for both the in-person and telephone efforts. The Craigslist postings were discontinued and a message on the recruitment website stated that the project was no longer registering volunteers. A final wave of recruiting was re-opened in May 2015 to augment the VO sample for the in-person effort in the Phoenix and Dallas CBSAs.

The Craigslist posting invited women 18 to 39 years old to participate in a study on health and safety (see Appendix D) and provided a link to a web page in English and Spanish where the volunteer could register their contact information after completing the eligibility screener (sex, age, and zip code). A deduplication process was implemented prior to registering the respondents to screen out recruits who signed up multiple times. Volunteers were offered a \$30 incentive to complete the interview. Respondents were not told about the actual topic of the survey until they were contacted by the interviewer to administer the survey.

The objective was to recruit approximately 2,000 women, ages 18-29, evenly distributed across the five CBSAs, and to randomly allocate recruits to each mode. If they were unwilling to complete the survey in the assigned mode, they were not permitted to participate in the study. The allocation between the modes was adjusted at several points during data collection to reflect the

different response rates achieved by mode to achieve the data collection targets of 1,000 completed VO interviews each for the in-person and telephone modes.

Volunteers who registered through the website were checked for age and sex eligibility as well as screened for duplicates based on phone number and first name. Women ages 18-29 were accepted with certainty. Volunteers ages 30-39 were informed they may be contacted if needed, but none were selected. After deduplication, records were assigned to either telephone or in-person data collection. Approximately 1.4 percent of the women ($n = 30$) in the 18-29 group who completed the main survey reported their age as 30 on the survey. This is likely due to turning 30 between registering on the website and taking the survey. Another 15 women reported an age older than 29 on the main survey. This may have been due to age entry errors or misrepresentation when registering. The VO respondents older than 29 were evenly distributed across the modes. (See Chapter 5 for actual age distribution.)

From September 30, 2014, through November 24, 2014, volunteers were divided equally across modes. Initial completion rates during this period indicated a higher completion of in-person interviews than by phone. In response, the proportion of volunteers was adjusted so that 70 percent of the volunteers processed from November 28, 2014, through January 8, 2015, were assigned to telephone data collection and 30 percent to the field. Sufficient volunteers were recruited to exceed the telephone target of 1,000 completed interviews with a final total of 1,136 completed VO interviews.

During the late winter and early spring of 2015, however, in-person VO interviews were not on pace to reach the 1,000 target. In response, the recruitment flyer was re-posted on Craigslist and 100 percent of the responses were assigned to the in-person mode from May 3, 2015, through May 28, 2015. As shown in table 4-6 below, the final allocation by mode for the entire VO recruitment period was 54 percent telephone and 46 percent in-person.

Table 4-6. Volunteer sample allocation by mode

Recruitment period	Telephone		In-Person		Total
	Sample	Pct.	Sample	Pct.	
9/30/14 to 11/24/14	1,510	51%	1,427	49%	2,937
11/28/14 to 1/8/15	779	71%	313	29%	1,092
5/4/15 to 5/28/15	-	0%	223	100%	223
Total	2,289	54%	1,963	46%	4,252

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Both modes had a similar completion rate for the VO sample (table 4-7). Telephone administration of the VO main interview ranged from a 45 percent completion rate in the Phoenix CBSA to 54 percent in Dallas. The overall telephone VO completion rate was 50 percent. The in-person administration of the VO main interview ranged from a 47 percent completion rate in the New York City and Dallas CBSAs to 60 percent in Los Angeles. The overall in-person VO completion rate was 52 percent.

Table 4-7. Volunteer sample results

	Phoenix	Los Angeles	Miami	New York City	Dallas	Total
Assigned to telephone mode						
Complete	165	261	159	320	231	1,136
Refusal	63	95	65	105	85	413
Other	138	175	119	193	115	740
Total	366	531	343	618	431	2,289
Completion rate	45%	49%	46%	52%	54%	50%
Assigned to In-person mode						
Complete	186	233	161	233	199	1,012
Refusal	44	39	25	61	36	205
Other	151	114	99	198	184	746
Total	381	386	285	492	419	1,963
Completion rate	49%	60%	56%	47%	47%	52%

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

4.3.2 Interviewing the Volunteer Sample

In September 2014, VO cases were released to the field so that supervisors were able to prepare for making assignments prior to the training held the following month. The release of VO cases occurred on a rolling basis throughout data collection as new volunteers signed up on the study website. Field interviewers made five contact attempts by telephone over the course of 10 days to reach the volunteer. These attempts included leaving voicemails and text messages. Field interviewers completed the same interview tasks as described above for the GP cases, excluding the household screening.

A total of 1,963 VO cases were released for in-person data collection. Of these, 1,012 completed the main interview, yielding a working response rate of 52 percent. Of the volunteers that

completed the main interview, 175 respondents were selected for re-interviews and 154 completed them.

The telephone interviews were conducted with the VO samples from October 2014 through early February 2015. Interviews were conducted as described above for the telephone GP sample (section 4.2). The VO sample was recruited on an ongoing basis during the data collection period as described above. A total of 2,315 VO cases were released for telephone data collection. Of these, 1,162 completed the main interview, yielding a working response rate of 50 percent.

4.4 Service Provider Sample

A separate set of respondents were recruited directly from agencies that provided support services to include individuals who were known to have been victims of rape and sexual assault. The original goal was to recruit individuals in large enough numbers to conduct quantitative analysis of their responses. As noted below, however, the number of individuals recruited did not meet these standards. Nonetheless, this group provided valuable feedback on how the questions worked and, especially, if the questions evoked particularly strong emotional reactions. These results are provided in Appendix K.

Women in the service provider (SP) sample were recruited with the assistance of agencies that agreed to distribute information about the study to their clients. Service providers were contacted in each of the CBSAs to solicit their participation. These contacts consisted of emails, phone calls, and brochures that explained the purpose and importance of the study, established the service provider's role in assisting with the study, and coordinated the logistics of recruitment with each agency. Agencies were identified for recruitment if they provided or were connected with longer-term services for women (such as counseling). Agencies that only provided short-term crisis services (such as sexual assault forensic examinations) were not selected for recruitment.

Service providers differed in their structure and resources so individualized participation plans were developed with each agency. These tasks included –

- distributing hard-copy study recruitment materials such as fliers, cards, and brochures
- talking with clients about the study
- emailing information about the study to the client list

- affixing their agency logo on study recruitment materials.

Outreach to service providers began early in 2014, and this outreach continued through February 2015. Actual recruitment of women from these agencies (i.e., the SP sample) began in early September 2014 and continued through September 2015. In all, 16 service provider agencies participated by distributing information about the study to their clients. The study recruited respondents from nine of these service providers with three agencies each in Dallas and in New York City and one agency each in Los Angeles, Miami, and Phoenix.

Recruitment of service provider agencies was a challenge. When agencies were reached, they expressed strong support of the study, but most did not actively participate. The agency recruitment process started with project staff conducting thorough online searches for service provider agencies in each MSA. The number of agencies available in each MSA varied widely, ranging from 3 in Miami to 34 in the New York City CBSA. Just over half of the recruited agencies (9 of 16) successfully provided clients who volunteered for the study.

4.4.1 Service Provider (SP) Respondent Recruitment

SP recruitment pamphlets and posters offered a \$30 incentive to SP respondents who completed the survey and, for in-person respondents, an additional \$10 to offset any travel expenses (e.g., to meet with the interviewer for an in-person interview). The materials provided individuals with a toll-free number and a study website address. The message on the toll-free number instructed callers to provide their state, contact information, and the name of the agency that referred them. Eligible respondents were allocated for the in-person or telephone mode of data collection. Table 4-8 provides the total number of participating agencies and volunteers.

Table 4-8. Service provider sample productivity

CBSA	Service provider agencies	SP responses
Phoenix	1	4
Los Angeles	1	17
Miami	1	6
New York City	3	15
Dallas	3	26
Total	9	68

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

SP sample cases were assigned to a select group of field and phone interviewers who had experience working with service provider agencies and/or interacting with trauma survivors. These interviewers received additional training regarding survivor issues and recognizing distress. Other than logistical issues regarding the location of the interview, the survey administration procedures were the same for all sample types (GP, VO, and SP).

4.4.2 Administration of the ACASI Interviews

Interviewers contacted SP cases as soon as they received the assignment and most appointments were set within one week of the case being assigned. There was very little non-response among SP cases and no refusals. Field interviewers completed the same interview tasks as described above for the VO cases. SP cases were not selected for re-interviews.

In-person interviews took place either at the service provider agency or a location designated by the respondent. For interviews that were conducted at the agency, service provider staff assisted in providing a safe and private location for the scheduled interview and, in some agencies, a counselor was available to respondents if needed during or after the interview. A total of 46 SP cases were sent for in-person data collection. Of these, 41 completed the main interview, yielding a working response rate of 89 percent.

4.4.3 Administration of the Telephone Interviews

Telephone interviews with the SP sample took place from October 2014 through early February 2015. From February through late May 2015, telephone survey operations were suspended for the RDD data collection. When the RDD phase began, a few additional SP sample interviews

were completed. SP telephone surveys were conducted in the same manner as the VO sample. A total of 22 SP cases were sent for telephone data collection. Of these, 17 completed the main interview, yielding a working response rate of 77 percent.

4.5 Field Interviewer Training and Monitoring

Field staff recruiting took place from June 2014 through September 2015. Hiring began with identification and selection of five experienced field supervisors to recruit field interviewers. The field recruiters contacted, screened, and interviewed field interviewer candidates. Only females were considered for the interviewer positions because of the sensitive nature of the survey. The hiring process yielded a total of 147 interviewers, 26 of whom did not complete training or were released from the study before receiving any case assignments.

4.5.1 Interviewer Training

Field interviewer training was administered in phases: (1) online home study, (2) in-person classroom training, and (3) post-classroom training activities. Each trainee completed a series of home study sessions prior to attending the in-person training. Classroom training sessions were held during the first 2 weeks of October 2014. Staggering the trainings across 2 weeks allowed the field manager to meet and become familiar with the field supervisors and field interviewers assigned to their CBSA. Each training lasted 5 days with supervisors receiving their portion of the training the first and last days (i.e., Monday and Friday) and the interviewer training take place on the three days in the middle. Sessions consisted of several modes of content delivery including scripted, video recording, and interactive hands-on exercises. The training also included a half-day session devoted to handling distress and special considerations for interviewing victims of sexual assault.

The supervisor training provided background on the study, procedures for case assignments, potential field interviewer challenges due to the sensitive nature of the survey, and other management roles/responsibilities. During the interviewer portion of the training, supervisors worked directly with their interviewers to provide instruction and lead exercises.

4.5.2 Field Interviewing Organization

The field interviewing management structure consisted of the director of survey operations, the field director, 3 field managers (FMs), 9 field supervisors (FSs), and 121 field interviewers (FIs). The five CBSAs were divided into regions consisting of one or more counties within the CBSA. This resulted in a total of nine regions across the five CBSAs. Each of the three FMs was responsible for monitoring all data collection activities in one or two CBSAs. One FS was assigned to a region within the CBSA and was responsible for the day-to-day supervision of all data collection activities within the region. FIs were assigned to a single region and a supervisor according to their geographical proximity to cases within the region. All FMs reported to the field director, who oversaw all aspects of data collection and communicated directly with the director of survey operations and other home office project staff. The field staff organization is shown in table 4-9 below.

Table 4-9. Field staff organization

CBSA	Field managers ^a	Field supervisors	Field interviewers
Phoenix	1	1	21
Los Angeles		1	24
Miami	1	2	19
Dallas		2	26
New York City	1	3	32
Total	3	9	122

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

^a There was one field manager for both Phoenix and Los Angeles, one for Miami and Dallas, and one for New York City.

Field supervisors were responsible for making all interviewing assignments. The general population assignments were made by geographic segments within a region. The caseload for a Field Interviewer depended on the size of the sample in a particular segment. In general, supervisors had a single interviewer work an entire segment in order to manage cases more efficiently. The average initial caseload per interviewer consisted of 60 cases.

During the field period, the supervisor monitored a set of daily productivity and performance metrics for each interviewer. These included the efficiency of interviewers in locating households in the GP sample, gaining their cooperation, completing the household roster, completing the interview process with eligible females, and, for a small number of cases, returning to the household to conduct a re-interview 2 weeks later. The interviewing tasks for the VO and SP sample consisted of arranging meeting places, making appointments for in-person data collection,

and conducting the in-person interview (and re-interview for selected VO cases). Contacting strategies for each case were dictated by the sample type.

Interviewers made in-person contact with an adult resident of the household to complete the roster task. An interviewer's GP assignments required them to visit multiple households within a segment, and interviewers developed weekly plans to ensure they worked efficiently and effectively. Work plans defined the days and times an interviewer planned to work, as well as the segments and number of cases the interviewer intended to work. It usually included details about refusal conversion strategies, plans for approaching locked buildings, and strategies for identifying non-viable cases (e.g., abandoned buildings, commercial buildings).

A Supervisor Management System (SMS) customized for this study provided field supervisors with several types of reports and systems to view groups of cases and individual case data. These included information at the case and task level. Case-level reports identified lists of cases that needed a particular action (e.g., the number of outstanding cases per interviewer). Task-level reports were used to identify specific tasks that needed attention (e.g., the number of outstanding re-interviews to be scheduled).

The field director and field managers generated additional reports nightly to identify specific trends and issues to be addressed. Managers met with their supervisors on a weekly basis to review overall data collection issues. Weekly meeting agendas included overall activity, numbers of contact, times/days of contact, and low productivity. Each week, the field director and field managers met with the survey operations director and other home office staff to provide updates and discuss data collection issues.

Two monetary incentive programs were established for in-person interviewers. The first bonus plan took place in December 2014 to boost productivity during the holidays. The second bonus program went into effect in March 2016 and continued to the end of the field period (June 2015). This bonus program incited interviewers remain on the study and boosted response rates for the GP Main Interview.

At the end of field data collection, debriefing sessions were held with interviewers and supervisors. Debriefings focused on sample types (GP-household, GP-college campus, VO, and SP) and the role of the field staff member. Questions were prepared and distributed to the field staff prior to the interview sessions. Responses from all of the debriefing sessions were compiled and reviewed with senior project management. (See Appendix D.)

4.6 Telephone Interviewer Training/Monitoring

The staffing process for the telephone data collection portion began in late summer to recruit and hire 18 interviewers for VO and SP data collection for the period from October 2014 to February 2015. The staffing requirement for the GP sample required 147 interviewers during the peak data collection effort. These interviews were conducted from December 2014 through early July 2015. Interviewers working during GP data collection included 15 who had worked on prior stages of this study. Due to the sensitive survey content, the telephone data collectors for this study needed to have a high skill level. Similar to the recruitment process for field interviewers, telephone data collectors for the SP sample underwent additional screening to ensure a very high level of sensitivity and professionalism.

4.6.1 Interviewer Training

The training for telephone interviewers was conducted over the course of a 6-day period and included 16 hours of self-paced learning, instructor-led video conferenced group sessions, and role-play in dyads. (See Appendix D.) Trainees were required to successfully complete each portion of training before they could proceed to the next portion. Those who had successfully completed the entire training protocol were sent information regarding scheduling hours for production and were eligible to begin working.

Once interviewers started working on production, additional training sessions were completed throughout the data collection period. Review sessions were held via video conference on a rolling basis after 2 weeks of production work and provided an opportunity for data collectors to talk about issues in gaining cooperation at both the household screening and extended interview levels, ask questions about unusual situations they may have encountered, and generally refresh their knowledge of study protocol and procedures. Team leaders maintained a webpage accessible only to telephone interviewers working on the study to post suggestions, updates regarding data collection progress, scheduling issues or changes, procedural clarifications, and reminders about specific aspects of the study protocol.

Interviewers assigned to special work classes (Spanish language interviews, re-interviews, and the SP recruited sample cases) received additional practice time to work on mock cases prior to

starting actual calls. A very small team of interviewers with high cooperation rates and sensitive study background/experience were assigned to the SP-recruited sample work class.

Interviewer feedback was shared with supervisors on an ad hoc basis and during the end of study debriefing sessions. Feedback indicated that the hiring and training approach was very successful in ensuring the desired level of sensitivity and professionalism this study warranted for front-line data collection staff. Some expressed that the pre-study interviews helped them feel like a real partner on the research team. Allowing all of the candidates an option to “opt out” after learning about the project and having any questions answered ensured that those who proceeded to the training sessions had a clear sense of the study content and sensitive nature before receiving case assignments.

4.6.2 Telephone Interviewing Organization

The telephone interviewing management structure consisted of three telephone data collection operations managers responsible for development of trainings and production management; two data collection project coordinators responsible for interviewer metrics, distress check-ins, and staff debriefings; 41 team leaders responsible for routine interviewer monitoring and coaching, shift scheduling, and taking inbound calls; and 147 telephone interviewers. Telephone interviewers had two forms of regular interaction with telephone center supervisory staff – routine monitoring and distress check-ins. On average, about 10 percent of interviewers’ scheduled interviewing time was monitored by telephone center team leaders (supervisors). Team leaders simultaneously connected to an interviewer’s phone conversations with potential respondents as well as viewed their computer screen for a 15-minute monitoring session, after which a monitoring sheet was completed. The interviewer’s performance in coding of results for call attempts, verbal interactions with respondents, and data entry of respondent answers to questions was assessed by the team leader.

Diversions from protocol were addressed on monitoring sheets and with follow-up telephone conversations with the interviewer as needed. Due to the relative rarity of interviews with respondents reporting past 12-month victimizations, team leaders were instructed to stay with a data collector conducting such an interview for a “double-length” monitoring session when they tap into such interviews. This was done to ensure that quality control efforts would encompass the extensive interview content asked only of persons with past 12-month victimizations.

After completing their first week of data collection, interviewers were contacted by a specific team leader assigned to the “distress check-in” role for a brief telephone meeting. These meetings provided interviewers with the chance to talk about any challenging interview situations and about their own feelings or issues around asking the sensitive survey questions. After this initial check-in, each interviewer continued to have a monthly call with the distress check-in team leader. For distress incidents that occurred during production interviews, a form completed by the interviewer describing the incident was immediately copied to the distress check-in team leader, who then contacted the interviewer for an ad hoc check-in. Feedback from the interviewers during project debriefing sessions at the end of the study indicated that the distress check-in process helped them feel confident and supported working on the project.

4.7 Field Problems – Difficulty Reaching Respondents

This section discusses issues that arose during the in-person data collection.

4.7.1 GP Sample

The procedures for contacting householders and respondents stipulated that field interviewers make five contact attempts over the course of 10 days. This included visits to the physical address as well as leaving an initial voicemail (when a phone number was provided) during one of the attempts and a follow-up text message. After five attempts, supervisors generally asked the interviewers to set these cases aside for a second-tier approach when and if resources allowed. In some cases, the supervisors made a telephone contact attempt and left a voicemail with their name and toll free number. Another second-tier strategy was waiting for a period of time to elapse before additional attempts. This approach was intended to extend the contacting period to anyone who may have had a period of time of being completely unreachable (e.g., an illness, being out of the country, having a baby, etc.).

Nonresponsive households posed considerable challenges for all interviewers. These included inaccessible dwelling units, refusals, deliberate avoidance, the inability to find residents of the household at home, broken appointments, and unwillingness to commit to an appointment for data collection. Field interviewers used numerous strategies to reach respondents and encourage participation. Their efforts and outcomes were documented in the digital records of contact. When

the field interviewer completed all contact attempts, as defined in the protocol, cases were transferred to their field supervisor for the assignment of a final disposition.

Inaccessible dwelling units were one of the biggest challenges for interviewers. These included locked buildings (e.g., apartments, condominiums) and gated communities. Throughout the data collection period, various strategies were used to facilitate contact, such as contacting property managers or security personnel to request access to the residential units, to hand deliver postcards. Field staff also used other strategies including mailing postcards and letters by U.S. mail and FedEx and looking up phone numbers for the sampled addresses. These efforts had marginal benefit.

4.7.2 VO and SP Sample

Volunteers provided their own phone numbers when registering for the study, so these were used instead of visiting a physical address. Otherwise, the procedures for contacting volunteers was similar to the GP sample. There were two main challenges inherent with VO cases: non-working phone numbers and reluctance to meet face-to-face to complete the interview. When interviewers encountered cases with non-working numbers, they tried after waiting 4 weeks. The reasoning behind this strategy is that the respondent's phone number may have been temporarily unavailable because of circumstances such as having an unpaid cell phone bill—thus, allowing additional time could result in a working number. In most cases, non-working numbers remained in this status and the case was assigned a final disposition. The interviewers encountered a higher proportion of non-working numbers early on the study. This may be due to the time elapsed between the initial Craigslist postings (late summer 2014) and the start of data collection (November 2014). The postings that brought in volunteers during the data collection period were more productive. These VO cases were released on a rolling basis and supervisors assigned them within 1 business day, and interviewers were asked to establish initial contact within 3 working days.

Some volunteers were hesitant to meet in-person with the field interviewer. Other reasons for refusal included respondents who wanted the interview to be administered over the telephone, were not interested or had forgotten about signing up for the study or had a change in life circumstances (for example, moved or got a job). The slightly higher response rate obtained for field versus telephone VO sample interviews is likely due to the use of text messaging as a contact protocol for field interviews. The text messages helped identify the contact as being part of the study for which the respondent had volunteered on Craigslist, and may have increased the likelihood of answering subsequent phone calls from the field interviewer.

Clients of service provider agencies also provided their own phone numbers and were the only group to be told during the recruitment stage that the study was about sexual victimization. As they were fully informed about the study purpose and had voluntarily signed up to participate, there was no difficulty with this sample in terms of refusals. However, this sample did experience unique challenges with scheduling interview times and locations. Several of the agencies had offered space in their office for the interviews so that their clients might be more comfortable answering questions about their victimization. While this was beneficial to the clients, this created logistical obstacles in some cases. Agency officers were not available at all hours and days of the week and interviewers had to negotiate appointment times that worked for both the client and the agency, which could be difficult. In addition, the interviewer had to coordinate with agency staff to ensure a private room was available at the appointment time. In some cases, agency staff were difficult to get ahold of on short notice. A second challenge was the fact that clients did not recognize the study phone number that was used by interviewers to call for appointments. Even though they had voluntarily signed up for the study, some clients did not make contact with the interviewers because they did not realize who was calling. To address this issue, interviewers later started texting clients after project staff had consulted with agencies on how to make contact with clients who did not answer their phones.

4.8 Telephone Problems – Difficulty Reaching Respondents

This section discusses issues that arose during the telephone data collection.

4.8.1 RDD Sample

The yield from the telephone survey for interviews with eligible female respondents was negatively impacted by a much lower than expected eligibility rate for the cell phone sample (which comprised over 80 percent of the entire RDD sample). For this sample, the estimated eligibility rate of reaching adult females at the screening level had been 40 percent. However, the observed eligibility rate was just 23 percent. This may be due to a difference in cell phone-answering behavior by sex. Interviewers assessed the sex of voice mail greetings and refusals and recorded whether the voices appeared to be male or female. Both indicated a higher rate of non-contacts and refusals by females compared to males. It is important to note that at the screening stage, the content of the main interview had not yet been revealed to potential respondents, which eliminates the possibility

that the higher rates of non-contact and refusals for females in the cell phone sample was driven by the survey content. This raises the possibility of a differential response by females to calls from unknown parties to one's cell phone.

An external change in federal requirements for conducting cell phone surveys also had a diminishing effect on the yield from the cell phone sample. In July 2015, during the RDD data collection, the FCC issued an update to the Telephone Consumer Protection Act (TCPA) of 1991 that tightened restrictions around the dialing of random cell phone samples. It was determined that some of the calling protocols should be revised in light of the restrictions. The changes that were implemented included –

- Cell sample phone numbers that had never been answered by a human across seven phone calls (always reaching voice mail or a ring no answer result) were finalized with no additional call attempts.
- Cell sample phone numbers undergoing refusal conversion attempts that had reached a maximum call total were finalized with no additional call attempts.
- No additional call attempts were made to refusals during which the respondent mentioned being on the “Do Not Call” list, asked to be taken off Westat’s list, or asked to never be called again.

The effect of these changes was a modest decrease in the RDD response rate for the sample released after the changes took effect in mid-August of 2015 (approximately 40 percent of the cell phone sample).

For the landline RDD sample, eligibility did not have a major impact on yield at the household screening stage. However, yield was negatively impacted by problems typical of landline samples in recent years—nonworking numbers (33% of landline screeners), phone numbers never answered by a person (38%), and refusals (16%). The landline sample screening stage response rate was just half that of the cell phone sample. The eligibility rate for reaching adult females in the landline sample was much higher (80%) than that obtained for the cell phone sample.

4.8.2 VO Sample

Telephone data collection with the VO sample obtained a 50 percent response rate, slightly lower than that obtained for the field VO interviews. Refusal to participate was the largest category of nonresponse to this effort, with 19 percent of the sample refusing the survey once contacted.

Volunteers refused the telephone survey for a variety of reasons, including those related to the sensitive survey content (which was explained verbally by the telephone interviewer upon making contact), and those related to the volunteer's life circumstances (e.g., some had volunteered for the financial incentive but in the interim found jobs and no longer needed the money, others simply were no longer interested in participating). As is typical with phone surveys, the other forms of nonresponse included phone numbers that were never answered after as many as 14 call attempts spread over the course of up to 3 weeks (14% of the sample), calls to numbers that were no longer working or at which those who answered claimed to have never heard of the volunteer (6% of the sample), and calls to numbers at which contact was made at some point but the survey was not completed before the calling algorithm maximum was reached (12% of the sample).

4.8.3 SP Sample

The telephone survey effort with SP sample respondents was much more successful than that with the VO sample. Interviews were successfully completed with 17 of the 22 SP sample cases assigned to the telephone survey (a 77% response rate). Nonresponse for the remaining five SP sample cases consisted of two cases that reached the maximum calling algorithm, two cases that never answered the phone (repeatedly reached voice mail), and one case that was not able to be reached at the phone number provided. None of the SP sample cases refused the survey request, likely due to the fact that the content of the survey was known at the time of volunteering for the research effort.

5. Response Rates

Chapter 3 described the sample designs for the RSA Pilot Test, emphasizing the importance of the number of completed interviews and their connection to the overall design goals for the study. The chapter discussed sampling variance as a measure of the precision of sample estimates and identified aspects of the sample design affecting the variance, including the completed sample size.

Response rates are an important factor in determining whether a study can meet its design goals for three reasons: First, planning the RSA Pilot Test, particularly the ACASI sample, involved estimating the initial sample size to ensure the target number of cases. This task required recognizing that the number of units selected would be reduced by both nonresponse and by the proportion of units without eligible respondents. Second, besides presenting challenges to achieving the target number of interviews, low response rates have an adverse effect on the cost per completed interview. Third, response rates continue to be generally regarded as a measure of the quality of survey outcome, although their importance has become a matter of dispute.

The first section of this chapter reports unweighted response rates for the GP samples following standards published by the American Association for Public Opinion Research (AAPOR). The second section further details the types of nonresponse affecting the ACASI and CATI samples. Because nonresponse may be correlated with survey design variables used to differentially sample the population, response rates are also frequently reported in weighted form. The third section then presents weighted rates, although the outcome is virtually identical to the unweighted analysis.

5.1 Overall Unweighted Response Rates

For a sampled respondent to be included in the RSA Pilot Test analysis, both the household screening and extended interview phases needed to be completed. For the ACASI and landline CATI samples, the household screening interview could be completed with any knowledgeable adult in the household. The screening interview was classified as complete if it provided sufficient information to determine which household members were eligible. At a minimum, determination of which female members of the household were age 18-49 was required for ACASI, and which female members were age 18 or over was required for the CATI landline sample. For the CATI cell sample,

the screening was complete if the respondent was a male, or if sufficient information on age was obtained from a female respondent to determine whether she was eligible.

As described in Section 3.3, both the sampled landline and cell phone numbers were groomed to remove unproductive numbers before being assigned to interviewers. These operations were performed by Marketing Systems Group's (MSG's) Comprehensive Screening Service (CSS) and Cell-WINS procedures, respectively. The sampled phone numbers excluded on this basis are omitted from the CATI nonresponse rates. An analogous exclusion arose for the ACASI sample, where cases were dropped if the USPS returned mail sent to their address as undeliverable on two occasions; these cases are omitted from the ACASI nonresponse rates.

AAPOR provides standards for reporting response rates (AAPOR, 2016) and an Excel implementation in a Response Rate Calculator V4.0 (downloaded Oct. 4, 2016, from <http://www.aapor.org/Communications/AAPOR-Journals/Standard-Definitions.aspx>). The standard describes an overall definition and guidance on implementation in specific applications. One of these applications is a dual-frame random digit dialing (RDD) survey, which must consider the overlap of the frames and the role of screening for eligible respondents in many applications. For each frame separately, the standard provides a general formula (AAPOR, 2016, p.69) that is, after simplification,

$$RR3 = \frac{I}{(I + P) + (R + NR + O) + [(UH)e_2]e_1 + [(UO)e_1]} \quad (5.1)$$

where

- $(I + P)$ = the number of complete and partial interviews,
- $(R + NR + O)$ = the number of refusals and other forms of nonresponse by eligible respondents,
- UH = the number of sample units with unknown household status,
- UO = the number of households with unknown eligibility status,
- e_1 = the estimated proportion of screener eligibility, and
- e_2 = the estimated proportion of household eligibility.

In applying the formula to the ACASI and CATI samples, no interviews were counted as partial, so $P = 0$. Other definitions for RR3 provided by the current AAPOR standards use only a single e rather than two, but for comparability formula (5.1) is applied for both ACASI and CATI.

The proportions e_1 and e_2 may be estimated from the data. The AAPOR standard suggests in this case

$$e_1 = \frac{(I + P) + (R + NR + O)}{(I + P) + (R + NR + O) + INR}$$

$$e_2 = \frac{(I + P) + (R + NR + O) + INR + UO}{(I + P) + (R + NR + O) + INR + UO + INNR}$$

where

INR =residential but ineligible for the survey, and
 $INNR$ =ineligible: not residential.

The denominator of e_2 includes all cases where the interviewer determined whether a household was present at the address or phone number, including the out of scope $INNR$ cases. The $INNR$ cases include those when the address was vacant, demolished, or temporarily occupied, such as a vacation home. In the case of CATI, the $INNR$ cases include those where the phone number was not working, belonged to a business, or had other nonresidential use.

A response rate for a survey that has to screen sample units for eligible respondents is often defined as the product of a response rate for household screening and a response rate for the extended interview. In fact, as will be illustrated here, the AAPOR formula (5.1) can be exactly expressed as this product when paired with specific definitions of the household screening response rate and the extended interview response rate. An additional notation, not included in the AAPOR standard, may improve the clarity of the underlying logic, namely, to recognize

$$IS = (I + P) + (R + NR + O) + INR$$

as the number of completed household screening interviews, which is the sum of the completed extended interviews, the number of noninterviews with eligible extended interview respondents, and the number of households determined not to include any eligible respondents.

The detail necessary to calculate RR3 is given in table 5-1, organized into the information used to compute a response rate for household screening and for the extended interview. The top line reports the initial sample of units. The second line reports the number of sample addresses

Table 5-1. Calculation of unweighted response rates by core-based statistical area for the ACASI general population sample

	Total	Phoenix	Los Angeles	Miami	New York City	Dallas
Total sample addresses ^a	22,249	4,787	4,094	4,425	4,700	4,243
Excluded USPS returns	862	385	34	280	57	106
Sample for household screening ^a	21,387	4,402	4,060	4,145	4,643	4,137
UH: Not contacted	1,014	34	547	120	91	222
INNR: Not a household	1,148	304	118	231	300	195
UO: Nonresponding households	8,542	1,457	1,586	1,506	2,186	1,807
IS: Household screening completed	10,683	2,607	1,809	2,288	2,066	1,913
e ₂ ^b	0.945	0.93	0.966	0.943	0.934	0.95
Household screening response rate ^{b,c}	0.529	0.637	0.461	0.586	0.476	0.487
Household screening completed	10,683	2,607	1,809	2,288	2,066	1,913
INR: No eligible respondent for extended interview	6,734	1,695	1,164	1,382	1,397	1,096
Eligible respondent for extended interview	3,949	912	645	906	669	817
(I+P) Completed extended interview	3,053	736	476	739	485	617
(R+NR+O): Extended interview nonresponse	896	176	169	167	184	200
Extended interview response rate ^b	0.77	0.81	0.74	0.82	0.73	0.76
e ₁ : Screening yield ^b	0.37	0.35	0.36	0.40	0.32	0.43
Overall response rate ^{b,c}	0.41	0.51	0.34	0.48	0.35	0.37

Note: Estimates are based on unweighted data. Response rate calculated using American Association for Public Opinion Research Response Rate 3 (RR3).

^aIncludes dormitory rooms for 10 out of 18 sampled colleges that cooperated with the survey but excludes addresses in one listed segment that were not processed.

^bThe rate for the total is computed as the unweighted average of the five CBSA rates.

^cThese rates do not reflect the unknown number of dormitory rooms in 8 sampled colleges that did not cooperate and the impact of the one segment that was not processed, out of 838 segments in the sample. The net overstatement of the response rates is approximately 1%; that is, the overall response rate is closer to 0.405 than 0.409.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

excluded from the nonresponse calculation because USPS returned mail twice as undeliverable. The unweighted response rate for the screening interview is based on the next lines of the table

Item in table	Explanation
Sample for household screening	Original sample with USPS returns removed
UH: Unknown status	Sampled units never assigned to an interviewer or not worked, unable to locate
INNR: Not a household	Vacant, demolished, commercial, etc.
UO: Nonresponding households	Evidence that household is present, but incomplete screening
IS: Household screening completed	Households successfully screened, whether or not they contain eligible subjects or complete an extended interview
e_2	The proportion of households among all sampled housing units after removing USPS returns

The unweighted screening response rate is computed according to the following formula:

$$\begin{aligned}
 RR3SCR &= \frac{(I + P) + (R + NR + O) + INR}{(I + P) + (R + NR + O) + INR + UO + e_2 UH} \\
 &= \frac{IS}{IS + UO + e_2 UH}
 \end{aligned} \tag{5.2}$$

The screening response rates in the five CBSAs approach or exceed 50 percent. Three CBSAs—Los Angeles, New York City, and Dallas—are similar, with values of 46.1 percent, 47.6 percent, and 48.7 percent, respectively. Phoenix and Miami are higher, at 63.6 percent and 58.6 percent. The average screening response rate is about 53 percent.

In (5.1) and (5.2), the factor e_2 is applied to UH , which represents sampled addresses where the presence of a household was undetermined. For the in-person sample, these are generally cases where no interviewer had time to contact the housing unit. The values of e_2 , which are close to 1.0, have the effect of counting most of UH as nonresponding households in (5.2).

The next section of the table 5-1 pertains to response to the extended interview. For clarity, the number of completed household screening interviews is repeated from a preceding line. The next lines of the table are

INR: No eligible respondent for extended interview	The screening interview determined that there was no eligible female ages 18-49 in the household
Eligible respondent for extended interview	Households with one or more respondents eligible for the extended interview
(I+P): Completed extended interview	Critical items of the extended interview are answered, allowing the case to be included in the analysis
(R+NR+O): Extended interview nonresponse	Any form of nonresponse for the extended interview, including refusals, respondent unavailable, and breakoffs

The response rate to the extended interview is calculated based on the eligible respondents. In this application the rate is defined as

$$RR3EXT = \frac{(I + P)}{(I + P) + (R + NR + O)} \quad (5.3)$$

Response to the extended interview ranges from 0.725 to 0.816, with an average of about 77 percent. On the next line, e_1 measures the yield from screening, that is, the proportion of screened households yielding an eligible respondent. The average is about 37 percent. Although not needed for the calculation of (5.3), e_1 appears in (5.1). Formula (5.1) is algebraically equivalent to the product of (5.2) and (5.3). The overall ACASI response rate is about 40 percent. The rates for the Phoenix and Miami CBSAs were notably better than for the remaining three CBSAs, as were the separate components of screener response rates and extended interview response rates.

Table 5-2 provides comparable nonresponse results for CATI landline sample, and table 5-3 presents comparable results for the cell sample. Rates for the extended CATI interview apply to the entire set of eligible respondents age 18 and over, rather than the 18-49 population. Unlike table 5-1, tables 5-2 and 5-3 do not report the number of cases excluded in advance, but the remaining lines have essentially the same interpretations as those in table 5-1. For both the landline and cell samples, a substantial proportion of the CATI samples remained with unknown status UH at the conclusion of the interviewing period. Most were phone numbers that when called produced a ring but no answer or only an answering machine or voicemail, without ever reaching a respondent.

Table 5-2. Calculation of unweighted response rates by core-based statistical area (CBSA) for the CATI landline sample

	Total	Phoenix	Los Angeles	Miami	New York City	Dallas
Sample for household screening	21,312	3,660	3,895	5,689	3,519	4,549
UH: Unknown status	8,032	1,693	1,477	1,719	1,308	1,835
INNR: Not a household	7,042	670	1,135	2,688	987	1,562
UO: Nonresponding households	4,072	784	879	882	794	733
IS: Household screening completed	2,166	513	404	400	430	419
e ₂ ^a	0.498	0.659	0.531	0.323	0.554	0.424
Household screening response rate ^a	0.213	0.213	0.195	0.218	0.221	0.217
Household screening completed	2,166	513	404	400	430	419
INR: No eligible respondent for extended interview	436	91	92	90	90	73
Eligible respondent for extended interview	1,730	422	312	310	340	346
(I+P) Completed extended interview	965	253	154	174	183	201
(R+NR+O) Extended interview nonresponse	765	169	158	136	157	145
Extended interview response rate ^a	0.56	0.60	0.50	0.56	0.54	0.58
e ₁ screening yield ^a	0.80	0.82	0.77	0.78	0.79	0.83
Overall landline response rate	0.12	0.13	0.10	0.12	0.12	0.13

Note: Estimates are based on unweighted data. Response rate calculated using AAPOR RR3.

^aThe rate for total is computed as the unweighted average of the five CBSA rates.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 5-3. Calculation of unweighted response rates by CBSA for the CATI cell sample and the combined overall rate for the CATI sample

	Total	Phoenix	Los Angeles	Miami	New York City	Dallas
Sample for household screening	105,498	19,746	23,562	24,112	18,927	19,151
UH: Unknown status	29,707	6,143	7,577	7,048	4,346	4,593
INNR: Not a household	8,953	1,765	1,903	2,110	1,697	1,478
UO: Nonresponding households	36,780	6,331	7,960	8,444	7,325	6,720
IS: Household screening completed	30,058	5,507	6,122	6,510	5,559	6,360
e ₂ ^a	0.882	0.870	0.881	0.876	0.884	0.898
Household screening response rate ^a	0.325	0.320	0.295	0.308	0.332	0.37
Household screening completed	30,058	5,507	6,122	6,510	5,559	6,360
INR: No eligible respondent for extended interview	23,128	4,211	4,878	5,020	4,274	4,745
Eligible respondent for extended interview	6,930	1,296	1,244	1,490	1,285	1,615
(I+P) Completed extended interview	4,222	854	747	858	764	999
(R+NR+O) Extended interview nonresponse	2,708	442	497	632	521	616
Extended interview response rate ^a	0.61	0.66	0.60	0.58	0.60	0.62
e ₁ screening yield ^a	0.23	0.24	0.20	0.23	0.23	0.25
Overall cell response rate ^a	0.20	0.21	0.18	0.18	0.20	0.23
Combined CATI response rate ^a	0.18	0.19	0.16	0.17	0.18	0.21

Note: Estimates are based on unweighted data. Response rate calculated using AAPOR RR3.

^aThe rate for total is computed as the unweighted average of the five CBSA rates.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

For the landline sample, the average overall response rate is 0.118, less than one-third of the ACASI rate. At 0.096, the rate appears distinctly lower in Los Angeles, with relative shortfalls in both the screener and extended interview rates. For cell phones, the average overall response rate is 0.198. Phoenix and Dallas show the highest response, with Los Angeles and Miami lower than the rest and New York City in the middle.

The screening yield is high for the landline sample, because a large proportion of households include one or more women age 18 or over. For the cell sample, the screening yield was about 25 percent or less, representing the combined effect of the sex and age restrictions.

Following AAPOR guidelines, table 5-3 provides a combined CATI response rate, which is an average of the landline and cell phone nonresponse rates, weighted in proportion to their completed sample sizes.

In summary, the screener response rates for ACASI sample in table 5-1 substantially exceed the corresponding rates for either the CATI landline or cell samples, both on average and in each CBSA. Similarly, response rates to the extended interview were also consistently higher in ACASI. For each mode, nonresponse to the screener questionnaire exceeded nonresponse to the extended interview; in other words, obtaining an initial contact and basic screening information was a greater obstacle than proceeding to conduct the extended interview after an eligible respondent had been selected.

5.2 Types of Nonresponse

Table 5-4 further details the sources of nonresponse for household screening. For the ACASI sample, sample addresses with unknown status were never visited or could not be located (previously reported as UH in table 5-1). Nonresponding households (UO in table 5-1) are divided into six categories in table 5-4. Refusals account for a significant proportion, although less than half, of household nonresponse. Three of the categories in table 5-4 represent circumstances where insufficient contact was made with the household: (1) reaching the maximum number of visits, (2) stopping work on the remaining sample at the end of the field period, and (3) being blocked from gaining access to the sampled units. Combined, these three categories represent the majority of household nonresponse cases in each CBSA. The number of households that were not accessible is a significant portion of the nonresponse, especially in Los Angeles, Miami, and New York City.

Table 5-4. Nonresponding households or persons to the screener by reason

	Total	Phoenix	Los Angeles	Miami	New York City	Dallas
ACASI						
Unknown status	1,014	34	547	120	91	222
Refusal	2,406	509	469	364	569	495
Maximum visits	2,475	635	344	321	584	591
End of field period	1,528	79	278	347	476	348
Unable to access	1,742	217	424	393	473	235
Language	143	10	58	15	40	20
Other	248	7	13	66	44	118
Screener cooperation rate	0.814	0.837	0.794	0.863	0.784	0.794
Landline sample						
Unknown status	8,032	1,693	1,477	1,719	1,308	1,835
Refusal	3,454	708	720	742	637	647
Maximum calls	295	47	52	62	90	44
Language	259	16	95	63	55	30
Other	64	13	12	15	12	12
Screener cooperation rate	0.385	0.420	0.359	0.350	0.403	0.393
Cell sample						
Unknown status	29,707	6,143	7,577	7,048	4,346	4,593
Refusal	30,153	5,460	6,348	6,648	5,946	5,751
Maximum calls	3,455	549	790	781	726	609
Language	3,087	306	804	1,001	634	342
Other	85	16	18	14	19	18
Screener cooperation rate	0.50	0.50	0.49	0.500	0.48	0.53

Note: Estimates are based on unweighted data.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The field staff tried several methods to gain access to the buildings or gated communities, including sending Federal Express letters to a subsample of the sampled cases. Nothing worked in a way that appreciably affected the ability to gain access. Language issues were cited in a fairly small number of cases. The residual category in the table, “Other,” is also small.

For the landline sample, the largest single contributor to the screening nonresponse rate is the set of sampled landline numbers with unknown status at the end of the calling period (UH in table 5-2), where it was not certain that the number belonged to a household. Of the nonresponding households (UO in table 5-2), the majority of them are refusals. There were a relatively small number of cases classified as reaching the maximum allowed number of calls, or as having language or other issues. For the cell sample, the number with unknown status at the end of the calling period is substantial but does not make up as much of the nonresponse as the landline sample, indicating more sampled persons are answering their phones.

Nonresponse rates combine the effect of noncooperation by respondents with difficulties of contacting the household and other obstacles to complete response. As an alternative measure of cooperation to reflect respondents’ actual willingness to participate once given the choice, the Response Rate Calculator V4.0 implements four definitions of the cooperation rate for a dual-frame RDD survey, which reduce to two distinct results when $P = 0$. For each RDD sample separately,

$$COOP4 = \frac{(I + P) + INR}{(I + P) + INR + R + UO} \quad (5.4)$$

The numerator of this rate counts both completed extended interviews and completed screening interviews that determined that no respondents were eligible. The results of this calculation appear in table 5-5. However, cooperation rates can be computed for the household screening and extended interview steps separately. The generic form of *COOP4* (AAPOR, 2016, p. 63) is simply,

$$COOP4 = \frac{(I + P)}{(I + P) + R} \quad (5.5)$$

Formula (5.5) is used in the last line of table 5-4 to produce screener cooperation rates for each CBSA separately, and the average of the five rates is shown in the last column. Screening cooperation was high for the ACASI sample, 78 percent or higher in each CBSA, and averaging about 81 percent. The cooperation rates are much higher and less variable among CBSAs than the

household screening response rates in table 5-1. In other words, non-contact and related obstacles to completing the screening interview are the principal sources of variation between CBSAs in the ACASI response rates.

For the landline sample, the range of screener cooperation varied fairly narrowly between a low of 35 percent and a high of 42 percent. The average, 38.5 percent, is almost double the average screening response rate, 21.3 percent, in table 5-2. Results for screening cooperation for the cell phone sample were somewhat higher than the landline sample, ranging from 48 percent to just over 52 percent. Cooperation rates are again higher than the screener response rates in table 5-3, which average 32.5 percent.

Response to the extended ACASI interview was high after screening, averaging almost 77 percent across the five CBSAs (table 5-1). Refusals accounted for between a third and a half of nonresponse to the extended interview (table 5-5). Non-contact with the sampled respondent contributed substantially as well, with little nonresponse due to language issues but somewhat more from the residual other category. The extended cooperation rate averaged 88.5 percent.

For the landline sample, response to the extended interview averaged about 55 percent across the five CBSAs (table 5-2). Refusals accounted for more than half of the extended interview nonresponse. The extended cooperation rate averaged about 67 percent.

Response averaged about 61 percent for the cell sample (table 5-3), with refusals again accounting for more than half of the extended interview nonresponse. The extended cooperation rate was 71 percent.

Table 5-5 also includes the overall cooperation rates based on (5.4). The averages are about 52 percent, 24 percent, and 42 percent, for the ACASI, landline, and cell samples, respectively.

5.3 Weighted Response Rates

When sample cases are selected with varying probabilities, weighted response rates can supplement the information from the unweighted versions by clarifying the effect of differential selection. In general, the final survey weights are not suitable for this purpose because they are only defined for the completed cases. Instead, base weights are available for all sample cases and are given in Chapter 3 by equations (3.2) and (3.6). They were used to form weighted response rates.

Table 5-5. Extended interview nonresponse by reason

	Total	Phoenix	Los Angeles	Miami	New York City	Dallas
ACASI						
Refusal	384	75	85	72	64	88
Non-contact	331	75	55	50	77	74
Language	31	5	6	7	8	5
Other	150	21	23	38	35	33
Extended cooperation rate	0.89	0.91	0.85	0.91	0.88	0.88
Overall cooperation rate	0.52	0.61	0.50	0.57	0.46	0.48
Landline sample						
Refusal	466	117	96	86	85	82
Maximum calls	129	27	16	22	38	26
Language	70	5	27	13	13	12
Other	99	19	19	15	21	25
Extended cooperation rate	0.67	0.68	0.62	0.67	0.68	0.71
Overall cooperation rate	0.24	0.28	0.20	0.21	0.24	0.25
Cell sample						
Refusal	1,722	300	306	388	313	415
Maximum calls	541	89	98	110	133	111
Language	252	27	57	85	37	46
Other	193	26	36	49	38	44
Extended cooperation rate	0.71	0.74	0.71	0.69	0.71	0.701
Overall cooperation rate	0.42	0.43	0.41	0.40	0.40	0.45

Note: Estimates are based on unweighted data.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The relatively minor weight variation for ACASI households has little effect on the response rates, as shown by the comparison of unweighted and weighted rates in table 5-6. The results agree closely. Because the CATI base weights for each telephone type within a CBSA are constant, the weighted response rates within each CBSA for the landline and cell samples are identical to the unweighted versions in tables 5-2 and 5-3. However, the overall response rate for the CATI sample can reflect the weighting, when it is formed as a weighted average of the landline and cell rates based on the weighted number of completed cases in the final sample. The weighted and unweighted rates are close, although the weighted rates are slightly less.

Table 5-6. Comparison of unweighted and weighted response rates by CBSA

	Total	Phoenix	Los Angeles	Miami	New York City	Dallas
ACASI unweighted rates						
Screener response rate	0.53	0.64	0.46	0.59	0.48	0.49
Extended interview response rate	0.77	0.81	0.74	0.82	0.72	0.76
Overall response rate	0.41	0.51	0.34	0.48	0.35	0.37
ACASI weighted rates						
Screener response rate	0.52	0.63	0.45	0.57	0.47	0.48
Extended interview response rate	0.77	0.81	0.73	0.82	0.72	0.75
Overall response rate	0.40	0.51	0.33	0.47	0.34	0.36
CATI overall response rates						
Unweighted	0.18	0.19	0.16	0.17	0.18	0.21
Weighted	0.18	0.19	0.16	0.16	0.17	0.21

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

6. Timing and Characteristics of Respondents

This chapter describes how long it took to complete the survey and the characteristics of the respondents. In both cases, these are described by mode and the two types of samples (general population; volunteer).

6.1 Time to Complete the Survey

Both instruments tracked the time that respondents spent overall and on different sections of the instrument (table 6-1). Timings are presented for the initial interview only. In order to reduce the influence of outliers, observations were capped for each section of the instrument at the upper 1 percent of the distribution separately by mode and sample type. Below, the results for the general population sample are described. The results for the volunteer sample displayed similar patterns.

Table 6-1. Average time to complete in minutes by survey section, sample type, and mode, 2014–2015

Survey section	Time to complete in minutes			
	General population		Volunteer sample	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
Household roster ^{c,d}	1.8	~	~	~
Extended interview overall ^e	25.6	25.4	31.3	31.2
Introduction and consent ^c	4.1 †	3.3	3.9 †	4.5
Demographics ^c	3.4 †	4.5	3.6 †	4.0
Event history calendar ^{c,f}	1.6	~	1.8	~
Sexual victimization screener	6.6 †	5.0	6.4 †	5.4
Detailed incident form	16.3 †	27.5	16.5 †	28.9
1st form	12.3 †	17.7	11.9 †	17.1
2nd form	7.1 †	12.2	6.6 †	11.7
3rd form	2.1 †	10.8	2.0 †	10.4
4th form	s	s	s	s
5th form	s	s	s	s
Vignettes	3.1 †	3.7	2.3 †	3.3
Respondent debriefing	1.9 †	2.4	1.4 †	2.1
Distress check-in ^c	0.4 †	1.2	0.6 †	1.2
Reinterview request/incentive ^c	2.4 †	2.8	2.6	2.7

Note: General population estimates are based on unweighted data for ages 18–49. Volunteer sample estimates are based on unweighted data for ages 18–29. See Appendix A for sample sizes and standard errors.

* Comparison group.

~ Not applicable.

† Significant difference from CATI at the 95% confidence level.

s Data suppressed for disclosure reasons.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cIn the ACASI condition, this section was administered by an interviewer using computer-assisted personal interviewing (CAPI), all other sections were self-administered using ACASI.

^dHousehold roster was administered to ACASI general population respondents only.

^eAverage timing starting with the introduction and consent. Does not include the household roster.

^fEvent history calendar was administered to respondents in the ACASI condition only.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

For the general population sample, the average time to complete the interview across modes was very similar. Overall, the average ACASI interview was completed in 25.6 minutes compared to 25.4 minutes for CATI. This difference on overall timing was not significant. However, the two modes did differ for several key sections, especially for those individuals who filled out a detailed incident form.

ACASI respondents took significantly longer to complete the introduction and consent section (4.1 vs. 3.3 minutes). This portion of the interview was self-administered for the ACASI respondents, whereas it was administered by the interviewer for the CATI. The ACASI respondents

spent an average of 1.6 minutes on the event history calendar section, which was not included in the CATI interview. The ACASI household roster took an average of 1.8 minutes for sample households, and this was also not included in the CATI interview.

The ACASI respondents took longer to fill out the sexual victimization screener than those on CATI. This may be because of the differences in the structure of this portion of the interview. The ACASI respondents were first asked about the occurrence of each type of victimization in the previous 12 months. If no past 12-month victimization was reported, they were asked about lifetime occurrence for the first eight screening items. The CATI screener asked about a lifetime event first. An affirmative answer led to asking about the previous 12 months. However, if no event was asked, the survey moved on to the next type of victimization. This difference led to asking more questions for the ACASI.

There were several sections where both surveys were administered by an interviewer. These included the demographics, the distress check-in, and the re-interview request. The interviewer for the ACASI sample did this in-person using computer-assisted personal interviewing (CAPI), while the CATI sample was administered over the telephone. In all three of these sections, the CATI took slightly longer to administer.

For those sections that were identical in structure, the self-administered mode took less time than the CATI. For the vignettes and respondent debriefing, the difference was relatively minor (e.g., 3.1 vs. 3.7 minutes for vignettes). However, there were larger differences in the amount of time it took to fill out a detailed incident form. On average, the CATI respondents that filled out at least one detailed incident form (DIF) spent more than 10 minutes longer doing so when compared to the ACASI respondents (27.5 vs. 16.3 minutes). The amount of time it took to fill out a form differed by which form was filled out. This was by design. Respondents were asked more questions for the first incident compared to the second and third one. For the first detailed incident form (DIF), it took ACASI respondents an average of 12.3 minutes compared to 17.7 minutes for the CATI. For the second DIF, it was 7.1 minutes for the ACASI and 12.3 for the CATI.

The ACASI interview took slightly longer than CATI when the respondent did not complete any DIFs (table 6-2). This can largely be attributed to the addition of the event history calendar portion of the interview in ACASI. The CATI interview took significantly longer than ACASI when any detailed incident forms were completed. For example, the average CATI interview when one DIF was filled out took 42.5 minutes compared to 34.6 minutes for the ACASI. CATI interviews

with two or more DIFs averaged a little over an hour. For ACASI it took 20 minutes less (44.9 minutes)

Table 6-2. Average time to complete extended interview by number of detailed incident forms, sample type, and mode, 2014–2015

Number of detailed incident forms	General population		Volunteer sample	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
Minutes	Minutes	Minutes	Minutes	Minutes
Zero	23.5 †	22.6	21.7 †	22.4
One	34.6 †	42.5	32.7 †	41.2
Two or more	44.9 †	65.4	43.2 †	62.2

Note: General population estimates are based on unweighted data for ages 18-49. Volunteer sample estimates are based on unweighted data for ages 18-29. See Appendix A for sample sizes and standard errors."

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The length of the DIF for the CATI is due to converting all questions with “yes/no” responses. This was done for confidentiality purposes so that anyone who might be overhearing the telephone call, for example in another room, would not be able to understand what was being asked. When questions had multiple response categories, this meant that all response categories had to be read aloud.

6.2 Demographics of Respondents

Prior to the victimization screener section, respondents were asked a series of demographic questions about their age, race, education, income, marital or romantic relationship status, employment status, military status, living situation, and public behaviors. This section presents the demographic characteristics of the respondents by sample type and mode.

There are two reasons for reviewing these data. The first is to provide the reader with a description of the types of individuals who are represented by the two different samples. While the general population sample was drawn to represent the five cities, it is restricted to a specific sex (female) and age group (18-49). The restriction to the five CBSAs also creates a subset that may differ from that of a national population. The second reason for reviewing these tabulations is to

provide information on whether the distribution of respondents differs between the modes for each sample type. Since many of the comparisons made throughout this report involve testing between modes, it is important to know if the distribution of respondents by these characteristics differs by mode.

For the tabulations below, the general population samples are based on the final survey weights. These weights were adjusted to the distribution in the targeted cities by age, race, marital status, and education for each mode of interview (Chapter 3). The distributions provided below for the volunteer sample are unweighted.

6.2.1 Race and Hispanic Origin

The majority of general population respondents were non-Hispanic white or Hispanic in both ACASI and CATI (table 6-3). The distributions across the modes are very close, as one would expect given that this was a variable used in the weighting.²¹ The largest difference is that ACASI had slightly more non-Hispanic whites (40.3% ACASI vs. 37.2% CATI) and Hispanics (36.9% ACASI vs. 36.2% CATI) compared to CATI. Slightly fewer general population ACASI respondents were non-Hispanic black (13.7% ACASI vs 14.0% CATI), another race (1.6% ACASI vs. 3.0% CATI), or reported more than one race (1.7% ACASI vs 3.1% CATI) compared to CATI. “Other types of races” reported included American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and individuals explicitly identifying as “other” race on CATI.

Overall, the majority of volunteer sample respondents in either mode reported that they were non-Hispanic white, non-Hispanic black, or Hispanic (table 6-4). In ACASI, 32.1 percent of volunteer sample respondents were non-Hispanic white, 25.7 percent were non-Hispanic black, and 32.1 percent were Hispanic. In CATI, 32 percent of respondents were non-Hispanic white, 23.3 percent were non-Hispanic black, and 31.7 percent were Hispanic. A larger percentage of the volunteer sample is non-Hispanic black when compared to the GP sample.

²¹The categories used for raking as part of the weighting were Hispanic, non-Hispanic black, and non-Hispanic other.

6.2.2 Age

There were no statistically significant age differences across mode for the general population (table 6-3).²² Overall, 62.6 percent of ACASI and CATI respondents were between the ages of 30 to 49. In ACASI, 8.9 percent of respondents were 18 to 20 years old, 12 percent were 21 to 24 years old, and 16.4 percent were 25 to 29 years old. Extremely similar distributions were found in CATI for these other age categories.

Table 6-3. Respondent demographics by mode of interview for females ages 18-49 in the general population, 2014-2015

	ACASI ^a	CATI ^b *
Race/Hispanic origin		
White ^c	40.3 %†	37.2 %
Black ^c	13.7 †	14.0
Asian ^c	5.8	6.5
Hispanic	36.9 †	36.2
Other ^{c,d}	1.6 †	3.0
More than one race ^c	1.7 †	3.1
Age		
18-20	8.9 %	8.2 %
21-24	12.0	12.9
25-29	16.4	16.3
30-39	31.3	30.9
40-49	31.3	31.7
Highest level of school completed		
No high school diploma	17.7 %†	13.4 %
High school graduate or GED	18.3 †	21.2
Some college, associate's degree, or vocational school	31.9	32.7
Bachelor's degree	22.8 †	20.0
Graduate degree	9.3 †	12.7
Currently enrolled in high school, college, or vocational school ^d		
Not currently attending or enrolled	79.3 %	78.4 %
College/university/graduate school	16.9 †	19.4
Member of a sorority	4.4 %	4.8 %
Other type of school	3.8 †	2.1

²²The categories used for raking as part of the weighting were 18-21, 22-24, 25-29, 30-34, 35-39, 40-44, 45-49.

Table 6-3. Respondent demographics by mode of interview for females ages 18-49 in the general population, 2014–2015 (continued)

	ACASI ^a	CATI ^{b*}
Total household income during 2013		
Less than \$25,000	31.3 %	32.3 %
\$25,000 – \$49,999	25.3	23.2
\$50,000 – \$75,000	13.9	13.3
More than \$75,000	29.5	31.2
Marital status		
Married	43.4 %†	42.4 %
Widowed	1.0	0.8
Divorced	9.2	8.5
Separated	4.0	5.0
Never married	42.4	43.2
Number of weighted sample cases	11,280,295	11,115,730

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cExcludes persons of Hispanic or Latina origin.

^dOther race includes American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and persons identifying as other race.

^eIncludes students currently enrolled part-time.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Similar to the general population, the volunteer sample did not have any mode differences on age (table 6-4). However, in comparison to the general population, the volunteer sample appears to be slightly younger. Most of the volunteer sample respondents were between the ages of 21 and 29 (82.3% ACASI vs. 80.6% CATI). In ACASI, 15 percent of volunteer sample respondents were 18 to 20 years old and 2.6 percent were 30 to 39 years old. In CATI, 17.8 percent of respondents were 18 to 20 years old, 1.4 percent were 30 to 39 years old, and less than 1 percent were 40 to 49 years old.

Table 6-4. Respondent demographics by mode of interview for females ages 18-29 in the volunteer sample, 2014-2015

	ACASI ^a	CATI ^{b*}
Race/Hispanic origin		
White ^c	32.1 %	32.0 %
Black ^c	25.7	23.3
Asian ^c	4.2	5.1
Hispanic	32.1	31.7
Other ^{c,d}	1.6	2.4
More than one race ^c	4.3	5.4
Age		
18-20	15.0 %	17.8 %
21-24	36.5	36.3
25-29	45.8	44.3
30-39	2.6	1.4
40-49	--	0.3
Highest level of school completed		
No high school diploma	9.1 %†	4.2 %
High school graduate or GED	18.1 †	21.8
Some college, associate's degree, or vocational school	45.5	47.0
Bachelor's degree	23.1	22.0
Graduate degree	4.3	5.0
Currently enrolled in high school, college, or vocational school ^e		
Not currently attending or enrolled	64.0 %	64.9 %
College/university/graduate school	31.9	32.4
Member of a sorority	3.5 %†	6.8 %
Other type of school	4.1	2.7
Total household income during 2013		
Less than \$25,000	46.1 %	43.6 %
\$25,000 – \$49,999	29.1	30.5
\$50,000 – \$75,000	11.4	10.9
More than \$75,000	13.4	15.0
Marital status		
Married	12.4 %	12.9 %
Widowed	0.2 !	0.1 !
Divorced	3.5	4.5
Separated	2.2	2.0
Never married	81.7	80.5
Number of unweighted sample cases	984	1,132

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

-- Less than 0.05%

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cExcludes persons of Hispanic or Latina origin.

^dOther race includes American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and persons identifying as other race.

^eIncludes students currently enrolled part-time.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

6.2.3 Education

Respondents were asked three questions related to educational attainment and school experience. First, respondents were asked to report their highest level of school completed or highest degree received. Second, respondents were asked if they are currently attending or enrolled in regular school such as a high school, or enrolled either full-time or part-time in college or university, trade, or vocational school. Third, respondents who indicated that they were enrolled in a college or university were asked if they are currently a member of a sorority.

In the general population, the distributions are similar to each other by mode (table 6-3).²³ There are slightly more respondents in the lowest education category (no high school diploma). In ACASI, 17.7 percent of general population respondents did not have a high school diploma, which was significantly higher than CATI respondents (13.4%). Slightly fewer ACASI respondents than CATI respondents were high school graduates or received a GED (18.3% ACASI vs 21.2% CATI). Slightly more ACASI respondents have a bachelor's degree than CATI respondents (22.8% ACASI vs 20% CATI). Finally, slightly fewer ACASI general population respondents have a graduate degree than CATI respondents (9.3% ACASI vs 12.7% CATI).

In regard to current school enrollment, the distributions are very similar between the modes. Most are not currently attending school (79.3% ACASI and 78.4% CATI). Slightly fewer ACASI respondents (16.9%) compared to CATI respondents (19.4%) are enrolled in college or university. Slightly more ACASI respondents are enrolled in another type of school (3.8%) compared to CATI (2.1%). Other types of school include high school, trade school, and vocational school. Out of those who are currently enrolled in college or university, 4.4 percent of ACASI respondents and 4.8 percent of CATI respondents are members of a sorority.

For the volunteer sample, there were more ACASI respondents with no high school diploma (9.1%) compared to CATI (4.2%) (table 6-4). Significantly fewer ACASI respondents had a high school diploma or GED (18.1%) compared to CATI (21.8%). In terms of current enrollment, there were no differences between the modes.

Across sample types, it appears that volunteer sample respondents have slightly higher levels of education than the general population. For example, 45.5 percent of volunteer sample

²³The categories used for raking as part of the weighting were: high school graduate or less, some college, bachelor of science or more.

respondents had some college, an associate's degree, or vocational school in ACASI and 47 percent in CATI. In terms of holding a bachelor's degree, 23.1 percent of ACASI respondents and 22 percent of CATI respondents had one.

The volunteer sample respondents also seem to be currently enrolled in school at a higher rate than the general population. For example, 31.9 percent of ACASI and 32.4 percent of CATI respondents are currently attending college or university. Of the respondents currently attending college or university, a significantly lower percentage of ACASI respondents are members of a sorority (3.5%) compared to CATI (6.8%). This is primarily because the volunteer sample is composed of women 18-29 years old, while the general population sample are in the 18-49 age range.

6.2.4 Income

Respondents were asked to report their total household income during 2013, before taxes, including income from work, investments, child support, and public assistance. Overall, there were no significant differences between the modes on income for the general population (table 6-3). Among ACASI respondents, 31.3 percent reported a household income of less than \$25,000, 25.3 percent reported an income of \$25,000 to \$49,999, 13.9 percent reported an income of \$50,000 to \$75,000, and 29.5 percent reported an income of more than \$75,000. The general population income distribution for CATI was extremely similar to ACASI's.

There are no significant differences between the modes on income for the volunteer sample (table 6-4). However, volunteer sample respondents seem to have a lower total household income than the general population. For example, 46.1 percent of ACASI respondents reported a household income of less than \$25,000 compared to 31.3 percent in the GP sample. The differences are likely related, in part, to the different age groups of the two samples.

6.2.5 Marital Status and Relationship Status

All respondents were asked about their current marital status, and those who were not currently married were asked if they are currently living with a romantic partner. Those who are not currently living with a romantic partner were asked if they have been in a romantic relationship in the last year.

Overall, most general population respondents reported that they were either now married (43.4% ACASI vs. 42.4% CATI) or have never been married (table 6-3) (42.4% ACASI vs. 43.2% CATI).²⁴ There were not meaningful differences among the marital status categories by mode of interview. The majority of general population respondents who are not currently married were not living with a romantic partner (table 6-5) (75.6 % ACASI vs. 76.8% CATI), and of those people, 54.5 percent in ACASI and 61.1 percent in CATI have been in a romantic relationship in the past year. While this difference is larger than all those observed, it is not statistically significant because this is a relatively small subgroup of those interviewed (i.e., those not currently married who are not living with a romantic partner).

Table 6-5. Respondent romantic relationship status by mode of interview for females ages 18-49 in the general population, 2014-2015

	ACASI ^a	CATI ^b *
Currently living with romantic partner?		
Yes	24.4 %	23.2 %
No	75.6	76.8
In a romantic relationship in the last year	54.5 †	61.1
Not in a romantic relationship in the last year	45.5 †	38.9
Number of weighted sample cases	6,388,544	6,396,884

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Contrary to the general population, the majority of volunteer sample respondents reported that they have never been married (table 6-4) (81.7% ACASI vs. 80.5 CATI). Only 12.4 percent of the ACASI respondents reported being currently married, as did 12.9 percent of the CATI respondents. Similar to the general population, the majority of volunteer sample respondents who are not currently married are not living with a romantic partner (table 6-6) (71.3% ACASI vs. 71.6% CATI), and of those people, 54.5 percent in ACASI and 61.1 percent in CATI have been in a romantic relationship in the past year.

²⁴The categories used for raking as part of the weighting were married and not married.

Table 6-6. Respondent romantic relationship status by mode of interview for females ages 18-29 in the volunteer sample, 2014-2015

	ACASI ^a	CATI ^b *
Currently living with romantic partner?		
Yes	28.7 %	28.4 %
No	71.3	71.6
In a romantic relationship in the last year	79.8	78.0
Not in a romantic relationship in the last year	20.2	22.0
Number of unweighted sample cases	861	986

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

6.2.6 Employment Status and Military Status

In the general population, significantly more ACASI respondents had a job or worked at a business last week compared to CATI respondents (64.9% ACASI vs. 58.6% CATI). In addition, table 6-7 shows that the majority of general population respondents reported that they never served on active duty in the military (97.6% ACASI vs. 97.4% CATI).

Similar to the general population, significantly more ACASI volunteer respondents had a job or worked at a business in the last week compared to CATI respondents (57.9% ACASI vs. 53.5% CATI). Table 6-8 shows that the majority of the volunteer sample respondents have never served on active duty in the military (96.1% ACASI vs. 97.5% CATI).

6.2.7 Home Ownership and Length of Stay

Respondents were asked if their home was owned or rented and how long they have lived at their current address. In the general population, 46.4 percent of ACASI respondents owned their home compared to 45.5 percent of CATI respondents, and 53.1 percent of ACASI respondents rent their home compared to 52.5 percent of CATI respondents. With respect to how long they have lived at their current address, significantly more ACASI respondents have lived at their current address for 1 year to less than 5 years compared to CATI (42% ACASI vs. 37.2% CATI). Significantly fewer ACASI respondents have lived at their current address for less than a year than

CATI respondents (17.9% ACASI vs. 23% CATI). About the same percentage of ACASI and CATI respondents lived at their address for more than 5 years.

Table 6-7. Respondent military status, employment, and living situation by mode of interview for females ages 18-49 in the general population, 2014-2015

	ACASI ^a	CATI ^{b*}
Had a job in the week before interview?		
Yes	64.9 %†	58.6 %
Served on active duty in the military?		
Yes, currently	0.2 %	0.2 %
Yes, in the past	2.2	2.4
No, never	97.6	97.4
Own or rent home		
Owned by you or someone in the household	46.4 %	45.5 %
Rented	53.1	52.5
Occupied without payment of rent	0.6 †	2.0
Length of time lived at current address		
Less than a year	17.9 %†	23.0 %
One year to less than 5 years	42.0 †	37.2
Five years or more	40.1	39.8
Number of weighted sample cases	11,280,295	11,115,730

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

In the volunteer sample (table 6-8), the majority of respondents rent their home (66.9% ACASI vs. 70.7% CATI). Significantly more ACASI respondents own their home compared to CATI respondents (30.9% ACASI vs. 25.2% CATI). Significantly fewer ACASI respondents occupy their home without payment of rent than CATI respondents (2.2% ACASI vs. 4.2% CATI). With respect to how long they have lived at their address, significantly fewer ACASI respondents have lived at their current address for less than a year (35.4% ACASI vs. 40.2% CATI), and significantly more ACASI respondents have lived at their home for 1 year to less than 5 years than CATI (39.1% ACASI vs. 32.3% CATI). Finally, in comparison to the general population, volunteer sample respondents appear to have lived at their current address for a shorter period of time as shown in tables 6-7 and 6-8. For example, 25.5 percent of ACASI and 27.5 percent of CATI volunteer respondents have lived at their home for 5 years or more. This compares to 40.1 and 39.8 percent for the general population ACASI and CATI samples, respectively.

Table 6-8. Respondent military status, employment, and living situation by mode of interview for females ages 18-29 in the volunteer sample, 2014-2015

	ACASI ^a	CATI ^{b*}
Had a job in the week before interview?		
Yes	57.9 %†	53.5 %
Served on active duty in the military?		
Yes, currently	0.6 %	0.2 %!
Yes, in the past	3.3	2.3
No, never	96.1	97.5
Own or rent home		
Owned by you or someone in the household	30.9 %†	25.2 %
Rented	66.9	70.7
Occupied without payment of rent	2.2 †	4.2
Length of time lived at current address		
Less than a year	35.4 %†	40.2 %
One year to less than 5 years	39.1 †	32.3
Five years or more	25.5	27.5
Number of unweighted sample cases	984	1,132

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

6.2.8 Public Behaviors

Respondents were asked about the frequency of three types of public behaviors in the last 12 months. First, respondents were asked how often they go shopping at drug stores, clothing, grocery, hardware, or convenience stores. Second, respondents were asked how often they spend the evening out, away from home, for work, school, or entertainment. Finally, respondents were asked how often they use public transportation. Respondents could describe their usage and frequency using the following response options: almost every day, at least once a week, at least once a month, less than once a month, and never.

In the general population, table 6-9 shows that the majority of respondents said that in the last 12 months, they had gone shopping at least once a week (60.5% ACASI vs. 61.6% CATI). In addition, slightly more ACASI respondents said that they shop almost every day compared to CATI

respondents (25.7% ACASI vs. 22.8% CATI). With respect to the frequency of spending the evening out away from home, there were no statistically significant differences between the modes for the general population. With respect to the frequency of public transportation use, slightly more general population respondents never use public transportation (60.3% ACASI vs. 58.1% CATI). Also, slightly fewer ACASI respondents used public transportation almost every day compared to CATI respondents (10.2% ACASI vs. 13.3% CATI).

Table 6-9. Respondent frequency of public behavior by mode of interview for females ages 18-49 in the general population, 2014-2015

	ACASI ^a	CATI ^{b*}
Average frequency of shopping last 12 months		
Almost every day	25.7 %†	22.8 %
At least once a week	60.5	61.6
At least once a month	10.8	12.7
Less often	2.6	2.6
Never	0.3	0.3
Average frequency spent the evening out		
Almost every day	19.4 %	19.6 %
At least once a week	35.9	35.7
At least once a month	20.6	21.6
Less often	16.0	14.0
Never	8.0	9.1
Average frequency of public transportation use		
Almost every day	10.2 %†	13.3 %
At least once a week	5.3	6.2
At least once a month	6.7	6.3
Less often	17.6	16.1
Never	60.3	58.1
Number of weighted sample cases	11,280,295	11,104,755

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

For the volunteer sample, table 6-10 shows that there were no mode differences for frequency of shopping. Most volunteer sample respondents reported that they go shopping at least once a week (57.2% ACASI vs. 58.6% CATI). When asked about the frequency that they spend the evening out away from home, more ACASI respondents reported that they spend the evening out at least once a week compared to CATI (41.7% ACASI vs. 36.1% CATI). In addition, fewer ACASI

respondents never spent the evening out compared to CATI (1.5% ACASI vs. 3.1% CATI). Finally, the volunteer sample did not have any mode differences in the frequency of public transportation use.

Table 6-10. Respondent frequency of public behavior by mode of interview for females ages 18-29 in the volunteer sample, 2014–2015

	ACASI ^a	CATI ^{b*}
Average frequency of shopping last 12 months		
Almost every day	29.7 %	26.3 %
At least once a week	57.2	58.6
At least once a month	11.0	12.9
Less often	1.8	1.9
Never	0.3 !	0.3
Average frequency spent the evening out		
Almost every day	28.2 %	30.3 %
At least once a week	41.7 †	36.1
At least once a month	18.1	19.3
Less often	10.6	11.1
Never	1.5 †	3.1
Average frequency of public transportation use		
Almost every day	26.4 %	28.0 %
At least once a week	9.0	9.1
At least once a month	8.1	9.3
Less often	19.0	17.6
Never	37.4	36.0
Number of unweighted sample cases	984	1,132

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

In summary, there is very close correspondence across modes for the above characteristics for both the general population and volunteer samples. For the general population sample, this indicates the weighting procedures were successful in at least standardizing the groups for these characteristics. For the volunteer sample, it indicates the random assignment process was successful in creating two groups that can be compared in the analyses described later in this report.

7. Prevalence and Incidence

This chapter provides the prevalence and incidence rates for the RSA Pilot Test and compares the rates to the NCVS and the National Intimate Partner and Sexual Violence Survey (NISVS).

7.1 Incident Classification Algorithm

To understand the types of sexual victimization experienced and the characteristics and outcomes of those experiences, the project team used the incident-based approach employed by the NCVS. This type of design asks RSA Pilot Test respondents to identify separate occurrences of victimization, date them, and then answer questions about each on a detailed incident form (DIF). Up to three DIFs were administered for each person.

The DIF is used to classify incidents into one of 40 types of crime categories. Two primary features of incidents were examined in order to classify incidents: (1) behavior and (2) tactic.

7.1.1 Behavior

The first part of the algorithm identifies the unwanted behavior that occurred during the incident. This includes the type of sexual contact that occurred, as well as whether it was threatened, attempted, or completed.

The following are the behaviors identified in this algorithm. Each type of behavior is classified in one of three categories – threatened, attempted, or completed:

1. **PENETRATION:** This includes all penetrative acts, including vaginal, anal, or oral. Behavior is classified as “penetration” if the respondent answers threatened, attempted or completed to any item D1A-D1D (table 7-1).
2. **SEXUAL CONTACT:** This includes all other types of sexual contact aside from penetration, including kissing, groping, exposure, or participation in sexual photos or videos. Behavior is classified as “sexual contact” if the respondent answers threatened, attempted or completed to any item D2A-D2D.

3. **MISSING BEHAVIOR:** An incident receives this code for behavior if the respondent replied “don’t know” or refused to answer at least one behavior item in D1A-D or D2A-D, and also said “no” to all other behavior items.
4. **NO BEHAVIOR:** An incident receives this code for behavior if the respondent answered “no” to all behavior items, with no refusals or “don’t know” responses.

These behaviors are listed in order of priority. For example, if a respondent indicates that an incident involved both penetration and sexual contact, then the incident behavior will be coded as “penetration.”

Table 7-1. Behaviors identified by detailed incident form

Behavior
Items identifying penetration
D1A Did the person threaten to, try to, or actually put his penis in your vagina when you didn't want it to happen?
D1B Did the person threaten to, try to, or actually put their mouth on your vagina or anus or make you put your mouth on their genitals or anus when you didn't want it to happen?
D1C Did the person threaten to, try to, or actually put his penis in your anus when you didn't want it to happen?
D1D Did the person threaten to, try to, or actually put fingers or another foreign object in your vagina when you didn't want it to happen?
Items identifying sexual contact
D2A Did the person threaten to, try to, or actually kiss or lick you when you didn't want it to happen?
D2B Did the person threaten to, try to, or actually touch, grab, or fondle your breasts, genitals, or buttocks over or under your clothes when you didn't want it to happen?
D2C Did the person threaten to, try to, or actually expose their sexual body parts or make you expose your sexual body parts when you didn't want it to happen?
D2D Did the person threaten to, try to, or actually make you look at or participate in sexual photos or movies when you didn't want it to happen?

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

7.1.2 Tactic

The second part of the algorithm identifies the tactic that was used to carry out the unwanted behavior.

The following are the tactics that are identified in this algorithm:

1. **FORCE:** This refers to physical force or threats of physical force. A tactic is classified as “force” if the respondent answers affirmatively to any item D4A-D4D (table 7-2).
2. **UNABLE TO CONSENT**²⁵: This is measured using three items asking about the effects of drugs or alcohol. A tactic is classified as “unable to consent” if a respondent indicates any of the following –
 - G10: The respondent was passed out for *at least part* of the incident.
 - G12A: The respondent said directly that the alcohol/drugs made her unable to give consent.
3. **COERCION:** An incident receives this code if the respondent answers affirmatively to any item D3B-D3D.
4. **TACTIC MISSING:** An incident receives this code for tactic if the respondent replied “don’t know” or refused to answer at least one tactic item listed above, and also said “no” to all other tactic items.
5. **UNWANTED:** This tactic indicates that the respondent answered “no” to all tactic items listed above, indicating no force, victim was able to consent, and no coercion, with no refusals or “don’t know” responses.

These tactics are listed in order of priority. For example, if a respondent indicates that an incident involved both force and inability to consent, then the incident tactic will be coded as “force.”

²⁵Also referred to “inability to consent” in the discussion below.

Table 7-2. Tactics Identified by detailed incident form

Tactic
Items identifying force
D4A Did the person ... hold you or pin you so you had difficulty moving?
D4B Did the person ... use a weapon, or threaten to use a weapon?
D4C Did the person ... physically attack, or threaten to attack you, but not with a weapon?
D4D Did the person ... physically attack, or threaten to attack someone else?
Items identifying inability to consent
G7 Do you think the person who did this to you was trying to get you drunk/high so they could sexually take advantage of you?
G10 Were you passed out for all or parts of this incident, or are you not sure? By passed out, it means that you were unconscious or asleep because of the alcohol/drugs.
G12A Did the alcohol/drugs make you unable to give consent?
Items identifying coercion
D3B Threaten to cut off financial support?
D3C Threaten to cause problems at your job, at school, in your relationships, or to cause some other problem?
D3D Promise rewards in your relationship, your job, your grades, or something else?

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

7.1.3 List of all Possible Classification Codes

By crossing the incident behaviors and tactics identified above, the algorithm classifies incidents into one of the following 40 incident types (table 7-3). These classifications are collapsed in various ways throughout this report in order to characterize the behaviors and tactics used in the incidents reported by respondents.

Table 7-3. Incident classification categories

	Behavior			Tactic	
	Sex	Contact	Force	Unable to consent	Coercion
	D1A – D1D	D2A – D2D	D4A – D4D	G7, G10, G12A	D3B-D3D
01. Completed forced penetration	Any = 3/Complete	Any response	Any=1/Yes	Any response	Any response
02. Attempted forced penetration	Any = 2/Attempt	Any response	Any=1/Yes	Any response	Any response
03. Threatened forced penetration	Any = 1/Threat	Any response	Any=1/Yes	Any response	Any response
04. Completed penetration while unable to consent	Any = 3/Complete	Any response	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any response
05. Attempted penetration while unable to consent	Any = 2/Attempt	Any response	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any response
06. Threatened penetration while unable to consent	Any = 1/Threat	Any response	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any response
07. Completed penetration using coercion	Any = 3/Complete	Any response	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any response
08. Attempted penetration using coercion	Any = 2/Attempt	Any response	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any response
09. Threatened penetration using coercion	Any = 1/Threat	Any response	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any response
10. Completed penetration with tactic missing	Any = 3/Complete	Any response	At least one=DK/RF AND Remainder=No		
11. Attempted penetration with tactic missing	Any = 2/Attempt	Any response	At least one=DK/RF AND Remainder=No		
12. Threatened penetration with tactic missing	Any = 1/Threat	Any response	At least one=DK/RF AND Remainder=No		
13. Completed unwanted penetration	Any = 3/Complete	Any response	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No
14. Attempted unwanted penetration	Any = 2/Attempt	Any response	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No
15. Threatened unwanted penetration	Any = 1/Threat	Any response	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No
16. Completed sexual contact using force	All=4/Didn't Happen	Any = 3/ Complete	Any=1/Yes	Any Response	Any Response

Table 7-3. Incident classification categories (continued)

	Behavior			Tactic	
	Sex	Contact	Force	Unable to consent	Coercion
	D1A – D1D	D2A – D2D	D4A – D4D	G7, G10, G12A	D3B-D3D
17. Attempted sexual contact using force	All=4/Didn't Happen	Any = 2/Attempt	Any=1/Yes	Any Response	Any Response
18. Threatened sexual contact using force	All=4/Didn't Happen	Any = 1/Threat	Any=1/Yes	Any Response	Any Response
19. Completed sexual contact while unable to consent	All=4/Didn't Happen	Any = 3/ Complete	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any Response
20. Attempted sexual contact while unable to consent	All=4/Didn't Happen	Any = 2/Attempt	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any Response
21. Threatened sexual contact while unable to consent	All=4/Didn't Happen	Any = 1/Threat	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any Response
22. Completed sexual contact using coercion	All=4/Didn't Happen	Any = 3/ Complete	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any=1/ Yes
23. Attempted sexual contact using coercion	All=4/Didn't Happen	Any = 2/Attempt	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any=1/ Yes
24. Threatened sexual contact using coercion	All=4/Didn't Happen	Any = 1/Threat	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any=1/Ye s
25. Completed sexual contact with missing tactic	All=4/Didn't Happen	Any = 3/Complete	At least one=DK/RF AND Remainder>No		
26. Attempted sexual contact with missing tactic	All=4/Didn't Happen	Any = 2/Attempt	At least one=DK/RF AND Remainder>No		
27. Threatened sexual contact with missing tactic	All=4/Didn't Happen	Any = 1/Threat	At least one=DK/RF AND Remainder>No		
28. Completed unwanted sexual contact	All=4/Didn't Happen	Any = 3/Complete	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No
29. Attempted unwanted sexual contact	All=4/Didn't Happen	Any = 2/Attempt	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No
30. Threatened unwanted sexual contact	All=4/Didn't Happen	Any = 1/Threat	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No

Table 7-3. Incident classification categories (continued)

	Behavior			Tactic
	Sex	Contact	Force	Unable to consent
	D1A – D1D	D2A – D2D	D4A – D4D	G7, G10, G12A
31. Missing behavior using force	At least one=DK/RF AND Remainder=4/Didn't Happen	Any=1/Yes	Any Response	Any Response
32. Missing behavior while unable to consent	At least one=DK/RF AND Remainder=4/Didn't Happen	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any Response
33. Missing behavior using coercion	At least one=DK/RF AND Remainder=4/Didn't Happen	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any=1/Yes
34. Missing behavior and missing tactic	At least one=DK/RF AND Remainder=4/Didn't Happen	At least one=DK/RF AND Remainder=No		
35. Missing unwanted behavior	At least one=DK/RF AND Remainder=4/Didn't Happen	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No
36. No behavior using force	All=4/Didn't Happen	Any=1/Yes	Any response	Any response
37. No behavior while unable to consent	All=4/Didn't Happen	All=2/No	G7=1 OR G10=1,2 OR G12A=1	Any response
38. No behavior using coercion	All=4/Didn't Happen	All=2/No	G7=2 AND G10=3,4 AND G12A=2	Any=1/Yes
39. No behavior with missing tactic	All=4/Didn't Happen	At least one=DK/RF AND Remainder=No		
40. No unwanted behavior	All=4/Didn't Happen	All=2/No	G7=2 AND G10=3,4 AND G12A=2	All=2/No

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

7.1.4 Narrative Examples of Classification Categories

Respondent narratives described the circumstances of the incident, including what happened during the incident, what the person did, whether there was a weapon, if alcohol or drugs were used, and any other details the respondent felt were important to understand what happened.

The examples in table 7-4 below are provided in order to illustrate some of the key DIF classification categories in table 7-3 based on the narrative responses. These come from both the volunteer and general population sample. The table lists several examples of narratives that were

classified as rape and sexual assault. Both of these involve either physical force or inability to consent. The examples involving coercion and unwanted acts are classified as ‘other unwanted sexual contact’ rather than rape and sexual assault.

Table 7-4. Examples of narratives for key DIF classification categories

Completed forced penetration	
Example 1	<i>"We went to my room and started kissing. He ... started taking my clothes off and we started having sex. ... He ... pinned me down with my arms behind my back. I yelled at him to stop ... but he put his penis in my anus. ..."</i>
Example 2	<i>"Heforced me to have sex even though I kept saying no. I had no way to leave and no phone to call someone for help."</i>
Attempted forced penetration	
Example 1	<i>"It occurred at a party where I was drunk and they were drunk and high. ...they started putting their hands in my..., holding me down. ... I kept struggling and protesting until ...they stopped long enough for me to break away."</i>
Example 2	<i>"He[got] on top of me. I told him to stop. He told me that I would like it.... He forced my legs apart ...He pulled down his pants and was pinning me by my throat.... I told him to stop...He let go and I got up and left."</i>
Completed penetration while unable to consent	
Example 1	<i>"I got drunk at a friend's house and passed out. While I was sleeping another guy came in the room and had sex with me. I woke up in the middle but ... passed back out...."</i>
Example 2	<i>"It was a party. ... We were all drinking...He put his mouth on my breast and fondled my butt and put his fingers in my vagina. I was sleeping the whole time."</i>
Attempted penetration while unable to consent	
Example 1	<i>"... He convinced me to take one shot with him ... , and don't remember much of what happened after that. My ... friend ... was able to stop him when he was trying to have sex with me ... even though I continuously tried to get away and say no."</i>
Example 2	<i>"I passed out. He came in and tried to kiss me. I told him no.... he was trying to take off my pants. ... I woke up. I told him no and then he got up and he left."</i>

Table 7-4. Examples of narratives for key DIF classification categories (continued)

Threatened forced penetration	
Example 1	<i>“...He tried to put his hands under my dress. I said no. He stopped ... but he was like whispering in my ear what he was going to do to me. ... When I tried to leave he wouldn’t let me go.”</i>
Example 2	<i>“.....a young man ... grabbed me and ... pinned me on the wall. He said he knew I wanted to ‘suck his penis.’ I ... shoved him away then walked fast somewhere public.”</i>
Completed penetration using coercion	
Example 1	<i>“He threatened to break up with me if I did not have what would be considered good sex for him....”</i>
Example 2	<i>“I saw my boss ... He kept trying to grope me. I said no ... As soon as we walked into the apartment, he threw me on the bed ... pinned me down ... Then from that evening on, he had threatened to fire me if I did not sleep with him. ...”</i>
Completed unwanted penetration	
Example 1	<i>“Sometimes my spouse wants to have sex when I don’t really want to do it. ...sometimes he keep insisting and bothering me until I consent.”</i>

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

7.1.5 Finalizing the Classification Scheme

Based on respondent narratives as well as other open-ended comments that provided additional detail about the incidents, some adjustments were made to refine and finalize the classification scheme. The project staff compared the narratives to the final classification to check if the incident was appropriately labeled. During this review process, the staff consulted with staff at AEquitas, a non-profit organization that provides prosecutors with support and training for cases involving gender-based violence and human trafficking. In addition, the staff met with the Office of Violence Against Women.²⁶

Narrative Review

In our consideration of the information provided in narratives, we reviewed the screening item that triggered the DIF, particularly incidents triggered by SV8 (attempted penetration against your will) and SV9-14 (various forms of sexual contact against your will). For each incident, two

²⁶The individuals consulted were lawyers in the Office on Violence Against Women, a component of the U.S. Department of Justice; and AEquitas, an organization providing resources for prosecutors of violence against women (<http://www.aequitasresource.org/>). These organizations bear no responsibility for the final decisions made on this project.

researchers re-read each narrative and independently coded whether the incident should remain in its current classification, or be down-coded or up-coded to another classification. We then adjudicated our coding with a third person.

SV8: Attempted Penetrative Acts. There was some concern that respondents may not have fully understood the meaning of an attempted act. Attempts can vary in interpretation as to what the perpetrator has to do in order to qualify. To qualify, the attempt had to be credible by what the offender did, as well as the context of the situation. For example, some respondents may interpret an individual verbally harassing someone when walking by on the street as an attempt. The victim may perceive this as a threat because of what the offender is yelling. However, this would not be considered a credible threat of penetration at that time. Consequently, we reviewed all incidents that screened in as an attempted penetrative act at SV8 and that were classified by the algorithm as an attempted act (n=121).

The narratives were not always complete and varied in quality. To use the narrative, we applied a rule that if it did not provide enough information or was ambiguous about the behavior or the tactic, we would rely on what was provided on the DIF to classify the incident. Only three incidents were changed. These included incidents that were clearly described as completed penetration rather than attempted penetration. For the 118 remaining narratives, the narrative supported the classification as attempted or there was not enough detail in the narrative to reclassify.

The table below shows a few examples of narratives for which the original classification was retained and one example for which the classification was changed (table 7-5).

Table 7-5. Examples of incidents involving attempted penetration in which classification was revisited based on narrative

Narrative	Original classification	Revised classification	Reasoning
"A gentleman I [had] sex with ... came to my address and tried to get through the door. ... threatened to knock the door down. I called the cops and before the cops had showed up, he had left."	Threatened forced penetration	Threatened forced penetration	Offender indicates intent and acts on it by coming to her home and trying to get in.
"I was at the person's house getting high. ... The guy tried to force himself on me and I was like telling him no and then I was like kicking at him and punching at him. Then I just ran."	Attempted forced penetration	Attempted forced penetration	Unclear what is meant by "the guy tried to force himself on me." We do not have enough details to merit changing the classification.
"I was at a ... party. ... I... couldn't walk straight or talk. That's when one of the males started to try and take me to the bedroom. I locked myself in ... but he found his way in."	Threatened penetration while unable to consent	Threatened penetration while unable to consent	While narrative does not provide details about what happened after he came into the room, there is not enough evidence to change the classification.
"I ... am not sure what they gave me to drink but I passed out and remember him on top of me ... having sex and I screamed no pushed him and he took me inside and pretended nothing happened."	Threatened forced penetration	Completed forced penetration	Respondent indicates the person was having sex with her.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

SV9-14: Unwanted Sexual Contact. We reviewed all incidents that screened in as unwanted sexual contact in items SV9-14 but that were classified by DIF responses as attempted or threatened penetration under force or inability to consent (n=81). We applied a rule in which all of these cases were down-coded to completed contact, and then reviewed each narrative in order to determine whether any of the DIF classifications should be coded to forced penetration or penetration while unable to consent (table 7-6). This change was based on two observations. One is that respondents seemed to be applying an overly broad interpretation of the "threat" or "attempt" question on the DIF. Our review of the narratives indicated that many respondents classified as threatened rape were responding to fears they had, but not necessarily something the offender did. Illustrations of this are described below. Second, these respondents had already said "no" to SV8, which specifically asked about an attempted rape.

We determined that 51 of 81 incidents could remain a completed contact. We developed several guidelines as we reviewed the narratives:

- If the person made threats to penetrate the woman in a public and populated space, like a bar or club, we did not consider this to be threatened rape because it was impractical to carry out the threat.
- If the person made such threats in a private setting or a public setting where no others were present, we considered this to be threatened rape.
- If the person was taking off, or trying to take off, the woman's clothing, or was groping her or kissing her unwantedly with no indication of intent to try to penetrate, we did not consider this to be attempted rape.

Table 7-6. Examples of incidents involving unwanted sexual contact in which classification was revisited based on narrative

Narrative	Original classification	Revised classification	Reasoning
<i>"I was walking through a crowd of people and a guy grabbed me around my waist and put his hand right on my butt. ... I told him no..., put my hand on his chest and pushed him away...."</i>	Threatened forced penetration	Completed forced sexual contact	No indication of any penetrative threat or attempt.
<i>"I ... walked by a table of guys and one of them grabbed my arm.... I pulled my arm away later, ... he grabbed ... my vagina, ... and said he wanted to try me out..."</i>	Threatened forced penetration	Completed forced sexual contact	Groping and suggestive language "he wanted to try me out" was made in a public place, but the threat was made in a public place and was not a credible threat.
<i>"He came in and started hugging me and fondling my boobs and my bottom. He was ...trying ...to take off my clothes. He was making sexual remarks and I wastrying to get him off of me. Finally I did and I left."</i>	Threatened forced penetration	Completed forced sexual contact	While the person did make sexual remarks while trying to take clothes off, there were no explicit actions toward penetrative behavior.

Table 7-6. Examples of incidents involving unwanted sexual contact in which classification was revisited based on narrative (continued)

Narrative	Original classification	Revised classification	Reasoning
<i>"...he immediately took off his pants and tried to have me do oraland he kept pushing my head down, and I kept saying no..."</i>	Attempted penetration while unable to consent	Attempted penetration while unable to consent	Even though this screened in as sexual contact, the narrative clearly indicates the person is trying to attempt oral sex.
<i>"We started making out andhe ... tried to put his hands in my vagina and anus.... I tried to get away and was pushing away from him..."</i>	Attempted forced penetration	Attempted forced penetration	Even though this screened in as sexual contact, the narrative clearly indicates the person is attempting digital penetration.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Use of Force

Use of force. The questions that were used to ask about tactics, such as force (D4a – D4d), included an open-ended item (D4e) asking about any other tactics that were not mentioned. In addition, if respondents answered “no” to all of the tactics in D4a-D4e, they were asked, “Was anything the person did that made this unwanted or against your will?” (D4f). If the answer to this was “yes,” the respondent was asked to describe the tactic. There were 179 incidents that used the open-ended response to one of these items to classify the incident as involving force. The types of tactics that were used to classify it as force included –

- held/pinned
- used weapon or threatened to use weapon
- physically attacked or threatened to attack, but not with a weapon
- physically attacked, or threatened to attack, someone else
- hit/punched/bit/choked/slapped/kicked/covered mouth
- grabbed/pushed/pulled/groped
- blocked/locked in/handcuffed
- other type of force.

The Person Stopped

As noted earlier, experts were asked to provide guidance on the criteria to use for the classification system. One issue of discussion was how to consider incidents in which the offender was engaging in unwanted behavior, but stopped immediately after the respondent said “no” or “stop.” The experts debated whether this might change an incident of attempted rape to unwanted behavior, thereby not being considered a crime. While our classification scheme is generally designed to comply with legal standards, this was a gray area, especially because it may be unclear if the person “immediately” stopped, or continued to attempt the behavior even though they may have stopped at one point or another. For purposes of classification, it was assumed that if the respondent said the person immediately stopped, we would not count this as force. As such, there were 151 incidents originally classified as forced sexual contact (i.e., touching or grabbing) or attempted/threatened forced penetration that were reclassified to unwanted contact, based on the respondent’s report that the person stopped immediately after she said “no” or “stop” (table 7-7).

Table 7-7. Examples of incidents when the person immediately stopped in which classification was revisited based on narrative

Narrative	Original classification	Revised classification	Reasoning
<i>“I was at a club.... As I continued to dance ... one of the guys... grabbed me, ... trying to kiss me. When ... I told him to stop, he ... let me go.”</i>	Attempted forced penetration	Completed forced sexual contact	Offender stopped when she asked him to.
<i>“He tried to have sex with me and fondle me but I refused, and he backed off and walked away....”</i>	Attempted forced penetration	Completed forced sexual contact	Offender backed off and walked away when she refused sex.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

7.2 Sexual Victimization Rates

This section presents estimates of incidence and prevalence rates using the detailed incident form for the general population. Appendix E provides similar rates using the SV screener and the VO sample. With the exceptions noted below, the results and conclusions are not different across the two methods by screening and type of sample.

7.2.1 Prevalence and Incidence of Sexual Victimization in the Past 12 Months

This section presents prevalence and incidence rates, both presented by mode.

Prevalence for the General Population for the past 12 Months

Prevalence rates were calculated by dividing the number of individuals victimized at least once by the total population.

Prevalence rates for the general population for rape or sexual assault within the last 12 months are 5.9 percent for ACASI and 5.3 percent for CATI (table 7-8). Incidents classified as rape or sexual assault are those where the respondent indicates they experienced unwanted sexual behavior involving force, threats of force, or while they were unable to consent due to drugs or alcohol. Incidents involving coercion or other tactics are classified as ‘other unwanted sexual contact’. To view a full table of prevalence rates for all 40 DIF classifications, see Appendix E.

Table 7-8. Prevalence rate of unwanted sexual contact, by type of crime classification and mode of interview for females ages 18-49 in the general population, 2014-2015

	ACASI ^a	CATI ^{b*}
Rape and sexual assault ^c	5.9 %	5.3 %
Completed ^d	4.4	4.2
Attempted ^e	1.8	1.5
Threatened ^f	0.7	0.4
Rape ^g	3.8 %	3.1 %
Completed ^d	2.4	2.1
Attempted ^e	1.2	1.0
Threatened ^f	0.6	0.4
Sexual assault ^h	2.8 %	2.7 %
Completed ^d	2.3	2.4
Attempted ^e	0.7	0.5
Threatened ^f	--	--
Other unwanted sexual contact ⁱ	8.1 %†	6.1 %
Completed ^d	5.3 †	4.1
Attempted ^e	3.1	2.3
Threatened ^f	1.0	0.8
Not enough information ^j	3.1 %†	1.4 %
No unwanted behavior ^k	2.4 †	1.3
Not enough behavior information to classify ^l	0.7 †	0.1 !
Not enough tactic information to classify ^m	0.2	--
Number of weighted sample cases	11,257,760	11,073,632

Note: Estimates are based on weighted data. See Appendix A for standard errors.

*Comparison group.

-- Less than 0.05%

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^dIncludes incidents where the offender completed the behavior.

^eIncludes incidents where the offender physically tried, but did not complete the behavior.

^fIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^gIncludes completed, attempted, threatened penetrative sexual contact using force or while unable to consent.

^hIncludes completed, attempted, threatened non-penetrative sexual contact using force or while unable to consent.

ⁱIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

^jIncludes incidents where respondent said "no," refused to answer, or didn't know the answer to all behavior and tactic items.

^kIncludes incidents where respondent said "no" to all items asking about unwanted penetrative and non-penetrative behaviors.,

^lIncludes incidents where respondent said "no," refused to answer, or didn't know the answer to all behavior items.

^mIncludes incidents where respondent said "no," refused to answer, or didn't know the answer to all tactic items.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The prevalence of rape is between 3 and 4 percent, depending on the mode (3.8% for ACASI, 3.1% for CATI), which includes incidents classified as attempted or threatened rape. The rates of completed rape are around 2.0 percent (2.4% for ACASI and 2.1 for CATI). The rates for attempted or threatened rape are about half of that (e.g., 1.2% for ACASI and 1.0% for CATI). The prevalence of sexual assault is 2.8 percent for ACASI and 2.7 percent for CATI. The rates for other unwanted contacts are much higher than either rape or sexual assault (8.1% for ACASI and 6.1% CATI).

The most frequently reported tactics related to rape and sexual assault incidents were those involving force, compared to those only involving inability to consent (table 7-9). For example, just under 2.0 percent of respondents indicated in the DIF that they experienced a completed forced rape (1.9% for ACASI, 1.7% for CATI), whereas only 0.5 percent reported a completed rape while unable to consent. The lower rates of incidents involving inability to consent partly reflect a hierarchy rule used in the classification. If the incident involved both tactics, it was put in the force category.

Table 7-9. Prevalence rate of rape and sexual assault by type of crime classification and mode of interview for females ages 18-49 in the general population, 2014-2015

	ACASI ^a	CATI ^{b*}
Rape and sexual assault ^c	5.9 %	5.3 %
Completed ^d	4.4 %	4.2 %
Forced ^e	3.8	3.7
Unable to consent ^f	0.7	0.6
Attempted ^g	1.8 %	1.5 %
Forced ^e	1.7	1.4
Unable to consent ^f	0.1 !	0.1 !
Threatened ^h	0.7 %	0.4 %
Forced ^e	0.6	0.2
Unable to consent ^f	0.1	0.2 !
Rape ⁱ	3.8 %	3.1 %
Completed ^d	2.4 %	2.1 %
Forced ^e	1.9	1.7
Unable to consent ^f	0.5	0.5
Attempted ^g	1.2 %	1.0 %
Forced ^e	1.2	0.9
Unable to consent ^f	--	0.1 !
Threatened ^h	0.6 %	0.4 %
Forced ^e	0.6	0.2
Unable to consent ^f	0.1	0.2 !

Table 7-9. Prevalence rate of rape and sexual assault by type of crime classification and mode of interview or females ages 18-49 in the general population, 2014–2015 (continued)

	ACASI ^a	CATI ^b *
Sexual assault	2.8 %	2.7 %
Completed ^d	2.3 %	2.4 %
Forced ^e	2.0	2.2
Unable to consent ^f	0.3	0.1 !
Attempted ^g	0.7 %	0.5 %
Forced ^e	0.6	0.5
Unable to consent ^f	0.1 !	--
Threatened ^h	-- %	-- %
Forced ^e	--	--
Unable to consent ^f	--	--
Number of weighted sample cases	11,257,760	11,073,632

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^dIncludes incidents where the offender completed the behavior.

^eIncludes holding or pinning, using a weapon or threatening to use a weapon, other physical attacks or threats of physical attacks on respondent or someone else.

^fIncludes incidents where respondents were passed out for all or parts of the incident or were unable to consent due to alcohol or drugs.

^gIncludes incidents where the offender physically tried, but did not complete the behavior.

^hIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

ⁱIncludes penetrative sexual contact using force or while unable to consent.

^jIncludes non-penetrative sexual contact using force or while unable to consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

For rape and sexual assault, there are no significant differences by mode, although the estimates for the ACASI are slightly higher than for the CATI. There are also no differences by mode when looking at incidents by tactics. There is a significant mode effect for other unwanted sexual contact. These are incidents involving unwanted behavior but no force or conditions related to inability to consent were involved. There is also a significant mode effect for the percentage of respondents who reported that no unwanted behavior occurred (2.4% for ACASI and 1.3% for CATI), although the magnitude of this effect is relatively small. These are instances where the respondent said “no” to all questions on the DIF that asked if any unwanted behavior occurred during the incident. This pattern is discussed in section 9.3.

To further explore the possibility of any differences by mode, a series of logistic regression models were estimated predicting the prevalence of rape and sexual assault using respondent characteristics, mode, and the interaction between mode and the characteristic. This further refines the comparisons made above by controlling for any differences in sample composition between the ACASI and CATI samples. It also tests whether there are any mode differences in the relationship between each characteristic by mode. The types of characteristics that were used in these regressions are shown in Appendix E. Examples include income, age, marital status and romantic involvement, involvement with school, and home ownership, among others. The results from these regressions are consistent with the tabular analysis discussed above. Only 2 of the possible 26 comparisons testing for a mode effect were statistically significant. Only five of the interaction terms were statistically significant. For more details on these results, see Appendix E.

Incidence Rate for the General Population for the Past 12 Months

Incidence rates were calculated by computing the average number of incidents experienced per 1,000 women. This accounts for instances where the individual was victimized more than once. These rates were calculated by counting the number of incidents for each type of victimization for each person. Because the respondent was not asked to fill out more than three DIFs, some incidents are not included in the incident rates. Respondents were asked to complete DIFs about the incidents involving penetration first, as reported in the SV screener.²⁷ Of the respondents who had at least one DIF classified as rape or sexual assault, approximately 9 percent (9.0% for ACASI, 8.9% for CATI) reported more than three victimizations in the past 12 months. See section 9.2 for a description of the frequency that respondents had incidents that were capped in this way.

The overall incidence rate for rape and sexual assault incidents per 1,000 women as measured by the detailed incident form is 84.9 for ACASI respondents and 77.0 for CATI respondents (table 7-10). To view a full table of incidence rates for all 40 classifications measured by the detailed incident form, see Appendix E.

²⁷Incidents reported in the SV screener were selected for a DIF using a selection algorithm that gave highest priority to forced rapes, then rapes while unable to consent, then unwanted penetration or attempted penetration, and then unwanted sexual contact/non-contact.

Table 7-10. Incidence rate of unwanted sexual contact per 1,000 females age 18-49, by type of crime classification and mode of interview in the general population, 2014–2015

	ACASI ^a	CATI ^{b*}
Rape and sexual assault ^c	84.9	77.0
Completed ^d	57.8	57.3
Attempted ^e	20.5	15.2
Threatened ^f	6.5	4.4
Rape ^g	51.0	43.1
Completed ^d	30.6	28.8
Attempted ^e	14.0	10.0
Threatened ^f	6.4	4.2
Sexual assault ^h	33.9	33.9
Completed ^d	27.3	28.5
Attempted ^e	6.5	5.2
Threatened ^f	--	--
Other unwanted sexual contact ⁱ	98.5 †	73.1
Completed ^d	55.9 †	41.9
Attempted ^e	31.8	23.5
Threatened ^f	10.7	7.7
Not enough information ^j	41.0 †	16.1
No unwanted behavior ^k	30.5 †	14.0
Not enough behavior information to classify ^l	8.8 †	1.7 !
Not enough tactic information to classify ^m	1.6	--
Number of weighted sample cases	11,280,295	11,115,730

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

-- Less than 0.05%

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^dIncludes incidents where the offender completed the behavior.

^eIncludes incidents where the offender physically tried, but did not complete the behavior.

^fIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^gIncludes penetrative sexual contact using force or while unable to consent.

^hIncludes non-penetrative sexual contact using force or while unable to consent.

ⁱIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

^jIncludes incidents where respondent said "no," refused to answer, or didn't know the answer to all behavior and tactic items.

^kIncludes incidents where respondent said "no" to all items asking about penetrative and non-penetrative behaviors, with no refusals or don't know responses.

^lIncludes incidents where respondent said "no," refused to answer, or didn't know the answer to all behavior items.

^mIncludes incidents where respondent said "no," refused to answer, or didn't know the answer to all tactic items.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Incidence rates for rape are 51.0 per 1,000 women for ACASI, compared to 43.1 per 1,000 women for CATI (table 7-10).²⁸ This is an average of 1.3 incidents for the victims identified for either mode. For completed rape, the incidence is 30.6 and 28.8 per 1,000 women for ACASI and CATI, respectively.²⁹ This is between 1.2 and 1.4 incidents per victim for the two modes.

As with the prevalence rates, the only significant mode effects are for the other unwanted sexual contact and no unwanted behavior.

Prevalence Rates for the Volunteer Sample

Up to this point, data have been presented for the general population sample (GP). In this section, the prevalence rates for the volunteer sample are presented because they exhibit a different pattern than the GP sample. As noted in the sample design section, the volunteer sample (VO) was considerably younger than the GP. This puts them in a higher risk group than the GP, and their victimization rates reflect this difference. Overall, 31.5 percent of ACASI respondents and 18.0 percent of CATI respondents in the volunteer sample reported experiencing at least one incident of rape or sexual assault within the past 12 months (table 7-11). Most of the incidents of rape and sexual assault were completed behaviors. The most frequently reported incident is completed rape (17.1% for ACASI respondents and 6.9% of CATI respondents). The rate of sexual assault is lower than for rape (11.1% ACASI, 9.9% CATI). The rate of other unwanted sexual contact was 32.5 percent for ACASI and 17.3 percent for CATI, which was the highest of all of the unwanted behaviors. There were 8.9 percent and 3.6 percent of the ACASI and CATI incident reports, respectively, that did not have any unwanted behaviors.

²⁸The priorities imposed on the administration of the DIF put force, threat of force, and inability to consent related to penetration at a higher priority than non-penetrative acts. This may contribute to why rape incidence estimates are higher than sexual assault estimates.

²⁹For example, the ratio between the incidence and prevalence for the ACASI completed rape estimates is 30.6/24, which is 1.27 to 1.

Table 7-11. Prevalence rate of unwanted sexual contact, by type of crime classification and mode of interview for females ages 18-29 in the volunteer sample, 2014–2015

	ACASI ^a	CATI ^b *
Rape and sexual assault ^c	31.5 %†	18.0 %
Completed ^d	25.2 †	13.3
Attempted ^e	8.9 †	6.5
Threatened ^f	3.3 †	1.2
Rape ^g	23.4 %†	10.3 %
Completed ^d	17.1 †	6.9
Attempted ^e	6.9 †	3.7
Threatened ^f	3.0 †	1.0
Sexual Assault ^h	11.1 %	9.9 %
Completed ^d	9.3	7.4
Attempted ^e	2.0	3.0
Threatened ^f	0.3 !	0.2 !
Other unwanted sexual contact ⁱ	32.5 %†	17.3 %
Completed ^d	21.5 †	10.7
Attempted ^e	10.5 †	6.5
Threatened ^f	6.4 †	2.6
Not enough information ^j	10.6 %†	4.0 %
No unwanted behavior ^k	8.9 †	3.6
Not enough behavior information to classify ^l	1.6 †	0.1 !
Not enough tactic information to classify ^m	0.9	0.4
Number of unweighted sample cases	983	1,124

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^dIncludes incidents where the offender completed the behavior.

^eIncludes incidents where the offender physically tried, but did not complete the behavior.

^fIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^gIncludes penetrative sexual contact using force or while unable to consent.

^hIncludes non-penetrative sexual contact using force or while unable to consent.

ⁱIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

^jIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all behavior and tactic items.

^kIncludes incidents where respondent said “no” to all items asking about penetrative and non-penetrative behaviors, with no refusals or don’t know responses.

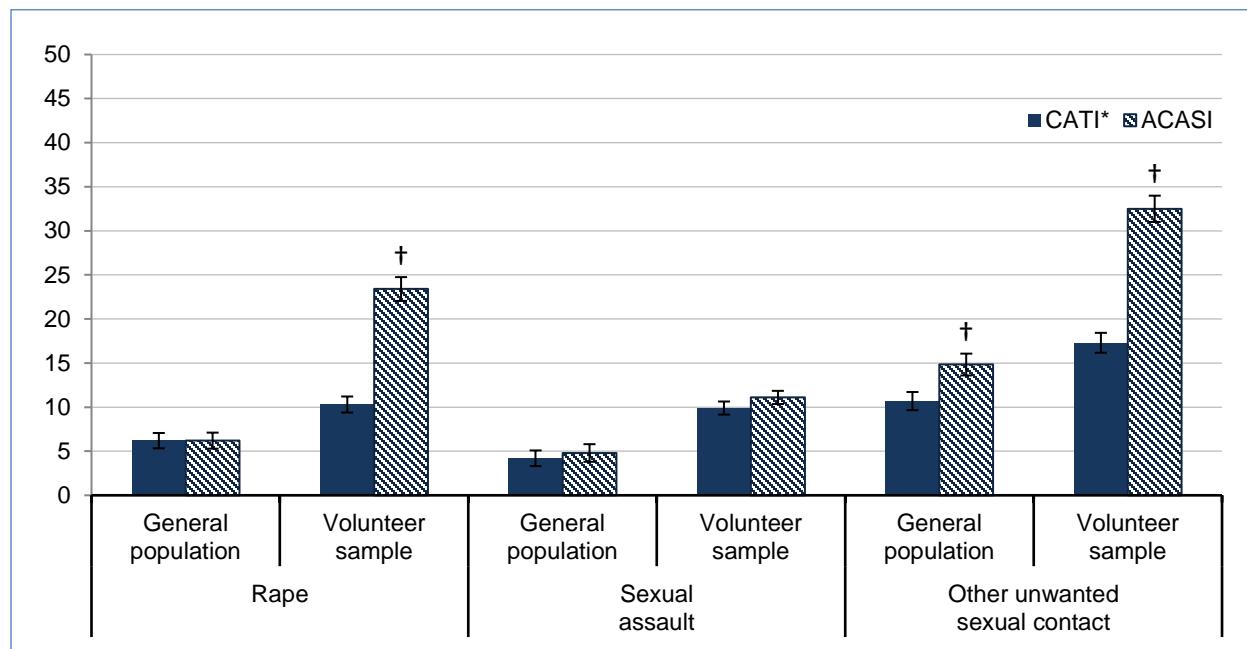
^lIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all behavior items.

^mIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all tactic items.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

With respect to mode, the VO sample is much different than the GP. There are large mode effects for rape and other unwanted sexual contact. For example, the prevalence rates of rape for ACASI are over twice as high as those from the CATI sample (23.4% vs. 10.3%). There is a comparable difference for the other unwanted sexual contact (32.5% for ACASI vs. 17.3% for CATI). No such mode effect was observed for the GP sample. These effects remain, even when restricting the GP sample to women who are age 18-29 (figure 7-1).

Figure 7-1. Comparison of mode effects in DIF prevalence estimates between general population and volunteer samples by victimization type for females age 18-29



* Comparison group.

† Significant difference from CATI at the 95% confidence level.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

To view volunteer sample prevalence rates for each of the 40 classifications measured in the detailed incident form, see Appendix E.

The stark differences by mode raises questions about the comparability of the VO sample by the two modes. As noted in the Data Collection chapter (Chapter 4), the VO sample was randomly assigned either to the CATI or ACASI condition over the course of the survey. Respondents were not given a choice but were assigned a mode. The assignment was checked in various ways, including comparing the personal characteristics, metro locations, and time period for those interviewed by ACASI and CATI. There were no substantive differences between the modes (see Appendix E). The two samples were interviewed over slightly different time periods. All of the

CATI respondents were interviewed between October of 2014 and January of 2015. For the ACASI, 80 percent of the interviews were completed during this time period. The other 20 percent were interviewed between February and May of 2015, with about half being done in May of 2015. Controlling for time period in the analysis of the victimization rates does not affect the significance of the mode effects discussed above.

7.2.2 Comparison to the NCVS and NISVS

Important benchmarks to evaluating the RSA Pilot Test estimates are comparisons to other surveys. In this section, the estimates are compared to the NCVS and the NISVS.

Comparison of Incidence and Prevalence Estimates to the NCVS

Incidence rates, per 1,000 females ages 18-49, were generated for the NCVS for the five CBSAs that were included in the RSA Pilot Test. In order to generate estimates with adequate statistical reliability, the NCVS rates were developed for the time period between 2011 and 2014 for the five CBSAs. The overall level of rape and sexual assault for this group of 18- to 49-year-old females was relatively stable over this time period, but there was an upturn in 2014. If these CBSAs followed the same trend, averaging over multiple years might produce an estimate that is slightly lower than what might have been observed for just 2014 if the sample sizes were sufficient.³⁰ Just as in the RSA Pilot Test, the NCVS rates for each CBSA were averaged together to derive the overall estimate shown in table 7-12.³¹

The RSA Pilot Test incidence estimates of rape and sexual assault are approximately 50 times higher than found for the NCVS. For example, the NCVS estimate for rape and sexual assault, the estimate typically published in the annual BJS reports, is 1.5 per 1,000 females ages 18 to 49. By comparison, the RSA Pilot Test has rates of 84.9 for the ACASI and 77.0 for the CATI (table 7-12). When breaking these out separately for rape and sexual assault, the differences are of similar magnitude. The separate estimates for rape and sexual assault for the NCVS are about the same (0.7

³⁰. Generated using the NCVS Victimization Analysis Tool at www.bjs.gov. 16-Oct-16. Bureau of Justice Statistics. Number and rates of rape and sexual assault, 1994 – 2014

³¹ The rate for each CBSA was calculated and an average of these rates was taken. This mirrors how the estimates for the RSA Pilot Test were calculated.

vs. 0.8).³² For both the ACASI and the CATI, the estimates for rape are slightly higher than for sexual assault, although these also have relatively large confidence intervals as well.

Table 7-12. Incidence rate per 1,000 of rape and sexual assault by study for females ages 18-49 in the general population

	NCVS ^a	Total	ACASI ^c	CATI ^d
Rape and sexual assault ^e	1.5	81.0 †	84.9 †	77.0 †
Rape ^f	0.7	47.1 †	51.0 †	43.1 †
Sexual assault ^g	0.8	33.9 †	33.9 †	33.9 †

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^aBased on NCVS for the years 2011-2014 among women ages 18-49 living in the 5 CBSAs of the RSA Pilot Test

^bBased on females ages 18-49 in the general population sample of the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^fIncludes penetrative sexual contact using force or while unable to consent.

^gIncludes non-penetrative sexual contact using force or while unable to consent. Note the NCVS estimate also includes unwanted sexual contact without force.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2011-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Prevalence estimates were also generated for the NCVS to compare to the RSA Pilot Test (table 7-13). The NCVS estimates are for females ages 18-49 living in CBSAs across the United States from 2011-2014. Note this is not restricted to the 5 CBSAs included in the Pilot. Generating NCVS prevalence rates for the 5 CBSAs was not operationally feasible. However, given the large differences between the RSA Pilot Test and NCVS estimates, expanding the NCVS data to include all CBSAs does not affect the conclusions below. Consistent with the incidence rates, the RSA Pilot Test estimates are considerably higher than those of the NCVS. The estimates for both rape and sexual assault are around 40 times higher (0.14% vs. 5.9% ACASI and 5.3% CATI for RSA Pilot Test). The rape estimates are also around 40 times higher (0.09% for NCVS vs. 3.8% and 3.1% for RSA Pilot Test), while the estimates for sexual assault are about 45 times higher.

³²These should be interpreted with caution because the confidence intervals for these estimates are very large (+/- .6).

Table 7-13. Comparison of prevalence rates from RSA Pilot Test to NCVS rates from 2011–2014 by type of victimization, study, and mode for females age 18–49 in the general population

	NCVS ^a	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Rape and sexual assault ^e	0.14 %	5.6 %†	5.9 %†	5.3 %†
Rape ^f	0.09 %	3.4 %†	3.8 %†	3.1 %†
Completed ^g	0.05 %	2.2 %†	2.4 %†	2.1 %†
Attempted ^h	0.03	1.1 †	1.2 †	1.0 †
Threatened ⁱ	0.01	0.5 †	0.6 †	0.4 †
Sexual assault ^j	0.06 %	2.8 %†	2.8 %†	2.7 %†
Completed ^g	0.05	2.3 †	2.3 †	2.4 †
Attempted ^h	~	0.6 †	0.7 †	0.5 †
Threatened ⁱ	0.01	--	--	--
Other unwanted sexual contact ^k	~ %	7.1 %	8.1 %	6.1 %

Note: Estimates are based on weighted data.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- NCVS estimate less than 0.005%. RSA Pilot Test estimate less than 0.05%

~ Not applicable.

^aBased on NCVS for the years 2011–2014 among women ages 18–49 living in all CBSAs in the United States.

^bBased on females ages 18–49 in the general population sample of the RSA Pilot Test, 2014–2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^fIncludes penetrative sexual contact using force or while unable to consent.

^gIncludes incidents where the offender completed the behavior.

^hIncludes incidents where the offender physically tried, but did not complete the behavior.

ⁱIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^jIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

^kIncludes penetrative or non-penetrative sexual contact where force or incapacitation was not reported. Note for NCVS, this category is included in completed sexual assault.

^lIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

Note for NCVS, this category is included in completed sexual assault.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2011–2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

There are several reasons why the incidence and prevalence rates for the two surveys differ. One is the scope of the incidents covered on the two surveys is different. The NCVS was last redesigned in 1992 when definitions and social response to rape and sexual assault were rapidly evolving. The NCVS victimization screener has two questions that ask specifically about rape and sexual assault. One of these is a probe that is embedded in an item asking about attacks and thefts.

This item specifically asks about “rape, attempted rape and other type of sexual attack” (Q41a on NCVS-1). The other question that targets sexual violence refers to “unwanted sexual activity” (Q43a on NCVS-1). In contrast, the RSA Pilot Test has 14 screening questions covering different behaviors (penetration, kissing, groping, exposure) and tactics (force; alcohol and drug use; unwanted). These provide the respondent with a wider definition of the target behaviors, as well as enhancing recall of incidents. The NCVS screener omits completely any items asking about a victim’s ability to consent. The RSA Pilot Test has a BSQ specifically targeting this tactic.

While the survey did not experimentally vary the number of screening items, it is possible to generate RSA Pilot Test estimates of rape and sexual assault once restricting it to incidents that are reported from screening and DIF items similar in scope to the NCVS. Initially, the RSA Pilot Test estimates were limited to –

- incidents generated by a ‘yes’ response to the screening items targeting completed and attempted forced rape (SV1-SV4, SV8)
- incidents generated by a ‘yes’ response to the screening items targeting completed and attempted sexual assault involving direct contact (SV9, SV10, SV11)
- incidents classified as a rape or sexual assault based on the DIF which involved physical force.

The first two restrictions are meant to simulate the types of incidents that are referenced on the NCVS screener. They eliminate from the RSA Pilot Test incidents that came from screener items that refer to behaviors that do not involve penetration (SV12, SV13, SV14) or other types of sexual contact or tactics that do not involve physical force (SV5, SV6, SV7). The second restriction takes out any incidents that were classified as RSA because of inability to consent because of alcohol or drugs. The NCVS DIF does not use this tactic in its operational definition. This filtering reduces the RSA Pilot Test incidence rate from 81 to around 62 per 1,000 women ages 18-49 (table 7-14).³³

A second type of screener item that is not on the NCVS is one that exclusively cues on attempted, but not completed, incidents. The NCVS screening instrument does include probes about attempts or threats at different points (e.g., Q40a). The probe in Q41a asks about “rape, attempted rape and other type of sexual attack.” There is no such mention or probe in the question targeting “unwanted sexual activity.” About 29 percent of the rapes on the RSA Pilot Test came from the question that specifically targets attempted unwanted sexual penetration and attempted

³³These estimates combine the ACASI and CATI responses.

sexual touching (SV8, SV11). If restricting estimates on the RSA Pilot Test once excluding the BSQ targeting attempts, the estimate drops to 35.

A third difference related to the scope of the incidents covered on the two surveys is the definitions emphasized on the victimization screeners. The NCVS does not ask about different forms of penetration, which make up the legal definitions of rape. As noted above, one screener item mentions the word “rape” in the probe. The ultimate classification of an event on the NCVS into the rape category relies on respondents identifying the incident as a rape when asked an open-ended question on the detailed incident form about how he/she was attacked (e.g., see Q27a, 28a, or 29a on NCVS-2). If the respondent does refer to the event as a rape or attempted rape, the interviewer confirms by asking, “You mentioned rape. Do you mean forced or coerced sexual intercourse?” Respondents may only be considering “sexual intercourse” to be vaginal penetration and, perhaps by just penile-vaginal penetration.³⁴

The reliance of relatively undefined terms, such as “rape,” “sexual assault,” and “unwanted sexual activity,” likely restricts the types of incidents that respondents think about. Many victims of rape do not generally think of the event in these terms. For example, on the RSA Pilot Test, respondents rarely used the word “rape” to describe what had happened to them even though their descriptions met the criteria defining it as rape. Restricting RSA Pilot Test estimates to just include the screening item that refers to vaginal penetration by a man’s penis using force or threats of force (SV1) and the sexual assault items (SV9, SV10), the rates are further reduced to 27 for rape and sexual assault.

In addition to differences in the relative scope of the incidents, a second design feature that may account for the differences is how the two surveys are framed for the respondent. The NCVS is introduced as a survey about crime. This may narrow the focus of NCVS respondents when thinking about eligible events. The title of the NCVS includes the word “crime.” The victimization screener begins with the following statement

“I’m going to read some examples that will give you an idea of the kinds of crimes this study covers.”

³⁴The NCVS Interviewer’s Manual explicitly defines sexual intercourse to include all types (digital, anal, oral), but the follow-up probe on the questionnaire does not define the term “intercourse” for the respondent. It is unclear how often an interviewer remembers to apply this rule, especially given the very rare instance that a respondent reports this type of incident. NCVS data editors do review the narrative provided for each incident. If the narrative conflicts with the classification based on the DIF, it is changed. Theoretically, this would catch instances of non-penile-vaginal penetration that may be classified as an assault.

The questions on rape and sexual assault are preceded by other types of predatory acts that are considered criminal (e.g., robbery, burglary, motor vehicle theft). In contrast, the RSA Pilot Test was introduced as a survey about health and safety. The questions do not refer to any acts as being criminal.

As will be discussed in Chapter 8, the survey included a question on whether the victim believed the incident was a crime at two different points in time—at the time of the incident and at the time of the survey. Restricting the RSA Pilot Test estimates to those incidents that the victim said it was a crime are further reduced to 7 and 15, depending on whether using opinions at the time of the incident or the survey (table 7-14). While still considerably higher than the estimates from the NCVS, this reduces the difference to a factor of 5 to 10, depending on the estimate and assumptions about what is a crime.

There are several other possible reasons why the RSA Pilot Test and NCVS estimates differ that are more difficult to quantify. One is the privacy of the interview. Everyone in the household age 12 and over is interviewed on the NCVS, and thus everyone knows what questions are asked on the survey. While NCVS interviewers are trained to try to conduct interviews in private, both in-person and telephone interviews may be conducted within earshot of other household members (Catalano, 2016). The RSA Pilot Test survey, in contrast, interviewed only one person per household and did not reveal the topic of the survey to any other member of the household. For the in-person visits, the interview was conducted using ACASI, so no other household member could hear any exchanges related to the topic or the survey items. Telephone survey respondents were encouraged to stop the interview if they thought someone in the household might be listening.

Another difference is a greater number of screening questions specifically asking about sexual victimization on the RSA Pilot Test compared to the NCVS (14 vs. 2). Survey methodologists have found that asking more questions that mention the targeted or related behaviors will produce higher rates of reporting. This was the basis of the redesign of the NCVS in 1992, which increased the measured victimization rates by increasing the density of the relevant cues (Biderman et al., 1986). Although the NCVS includes a number of screeners pertaining to general criminal victimization, exposure to multiple questions on different types of sexual violence on the RSA Pilot Test could have triggered memory of incidents that would not have otherwise been reported. Six of the 14 BSQs on the RSA Pilot Test were selected to enumerate specific behaviors and tactics that constitute rape. These questions elicited incidents that accounted for about 80 percent of the rapes reported on the survey. The remaining 20 percent were reported at BSQs that asked about other types of behaviors and tactics. Approximately 75 percent of sexual assaults were reported at three

different items targeting non-penetrative sexual contact and non-contact and 25 percent were reported at screener items that mentioned a different behavior. As noted above, the NCVS includes two questions that directly ask about rape and sexual assault.

One other difference that cannot be quantified is that the RSA Pilot Test is a one-time survey conducted by a private contractor, while the NCVS is a rotating panel design administered by the U.S. Census Bureau. The NCVS panel design would be expected to lead to lower victimization rates for two reasons. One is because interviews are bounded and are subject to panel conditioning (e.g., see section 9.1). On the NCVS, for example, violent crime rates at the second time in-sample are twice as high as the seventh time in-sample (Couzens et al., 2014).³⁵ Due to telescoping and conditioning, the NCVS downweights the first interview by factors as low as .5. Finally, the NCVS has a significantly higher response rate than the RSA Pilot Test. In the particular cities that were included in the RSA Pilot Test, the average response rate on the NCVS was approximately 28 percentage points higher than the ACASI survey and 50 percentage points higher than the CATI survey. If non-respondents are more likely to be non-victims, this would lead to artificially inflating the RSA Pilot Test estimate.

The study did not experimentally vary these design components to assess their effects on the estimates. However, the conclusions associated with the screener and context are consistent with a quasi-experiment, which compared the use of BSQs and a non-crime context to a survey that replicated the NCVS methodology (Fisher, 2009). In this study, both surveys were completed by telephone. The survey using BSQs also had a tailored DIF similar to what was used on the RSA Pilot Test. The survey using the BSQs was introduced as a survey on sexual victimization, rather than on criminal victimization as with the NCVS. Consequently, the comparison between the two reflects both the screening and context effects discussed above. The difference between the rates of rape for college students was a factor of 10 (44.8 vs. 4.5). A similar comparison from table 7-14 for the RSA Pilot Test ranges from 5.5 (81 vs 14.7) to 11 (81 vs. 7.3).

³⁵Many attribute this drop to “respondent fatigue.” But it may also be due to respondents’ deeper understanding of the requirements of the survey (e.g., Cantor, 1989; Kroh et al., 2016).

Table 7-14. Incidence rates of rape and sexual assault per 1,000 females age 18-49 by victimization screener item generating the report and whether respondent believed it was a crime, in the general population, 2014-2015

	Respondent considered it a crime at—		
	All incidents ^a	Time of incident ^b	Time of survey ^c
RSA Pilot Test – Incidents generated by all 14 screener items ^d	81.0	25.7	48.2
RSA Pilot Test – Incidents generated by 8 screener items: ^e	61.7	20.0	37.4
– Completed penetration by physical force and attempted penetration			
– Completed and attempted non-penetrative sexual contact against will			
RSA Pilot Test – Incidents generated by 6 screener items: ^f	34.8	11.2	21.2
– Completed penetration by physical force			
– Other completed non-penetrative sexual contact against will			
RSA Pilot Test – Incidents generated by 3 screener items : ^g	27.0	7.3	14.7
– Completed penile-vaginal penetration by physical force			
– Other non-penetrative sexual contact by physical force			
NCVS rape and sexual assault ^h	1.5	~	~

Note: Estimates are based on weighted data. See Appendix A for standard errors.

~Not applicable.

^aIncludes all rape and sexual assault incidents regardless of whether respondent said that they considered the incident to be a crime.

^bIncludes rape and sexual assault incidents where the respondent said that at the time of the incident they considered it to be a crime.

^cIncludes rape and sexual assault incidents where the respondent said that at the time of the survey they considered the incident to be a crime.

^dIncludes incidents classified by the DIF as rape and sexual assault.

^eIncludes incidents classified by the DIF as rape and sexual assault generated by screener items SV1–SV4, SV8-SV11. Excludes incidents classified by the DIF as rape or sexual assault because of an inability to consent.

^fIncludes incidents classified by the DIF as rape and sexual assault generated by screener items SV1-SV4, SV9, SV10. Excludes incidents classified by the DIF as rape and sexual assault because of an inability to consent.

^gIncludes incidents classified as rape and sexual assault generated from screener items SV1, SV9, SV10. Excludes incidents classified as rape or sexual assault because of an inability to consent.

^hBased on NCVS data for 2011-2014 among women ages 18-49 living in the 5 CBSAs of the RSA Pilot Test .

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Comparison of Prevalence Estimates to the NISVS

This section compares sexual victimization prevalence rates from the RSA Pilot Test to the most recent publicly available National Intimate Partner and Sexual Violence Survey (NISVS) estimates. The NISVS is conducted by the Centers for Disease Control and Prevention (CDC) and uses behaviorally specific questions to measure rape and other types of sexual violence. As discussed

in Chapter 2, several of the questions on the RSA victimization screener are based on items from the NISVS. However, there are several key differences between the RSA Pilot Test and the NISVS. One difference is the context of the questions. The NISVS asks about a wide range of sexual violence and other types of victimization (e.g., stalking), with some of these questions preceding the rape items. Within the sexual violence section, the order of the questions is different. Non-penetrative sexual contact is asked first, then alcohol/drug-related penetration, and then penetration related to force. This is opposite of the RSA Pilot Test, which first asked about forced penetration, then alcohol/drug-related penetration, other types of unwanted penetration, and then non-penetration sexual contact and non-contact.

Finally, because of sample sizes, the NISVS estimates for this section were generated for the entire United States, which differs from the RSA Pilot Test estimates, which are for specific cities.

The NISVS estimates use responses to the behaviorally specific questions to estimate rape and sexual assault. For purposes of comparison, two sets of estimates are provided from the RSA Pilot Test. One is from the SV screener items. For this purpose, SV screener items 1-5 were used to generate the estimates of completed rape. Methodologically, this is the most comparable estimate to NISVS. The second estimate is from the DIF.

Completed Rape. Twelve-month prevalence estimates of completed rape are similar between NISVS and the RSA Pilot Test. The NISVS estimates are nominally lower than the RSA Pilot Test's for completed rape in the past 12 months, but only one is statistically different. For the NISVS, the rate is 1.8 percent, which is within one percentage point of all four RSA Pilot Test estimates (table 7-15). The estimates are closest when comparing the CATI estimates, although one of the four possible RSA Pilot Test estimates is statistically different from the NISVS (SV from the ACASI).

There is a larger difference between studies when looking at completed rape by tactic. Three out of the four RSA Pilot Test estimates of completed forced rape were about twice as high as the estimates coming from NISVS (1.7-1.9% for RSA, 0.8% for NISVS). Conversely, for completed alcohol/ drug-facilitated penetration, rates from the two RSA Pilot Test DIF estimates were less than half the rate estimated by NISVS (0.5% for RSA DIF, 1.2% for NISVS). The RSA Pilot Test for alcohol/drug-facilitated penetration using the SV classification resulted in estimates similar to those seen in NISVS. These differences are related to the ways the two methods classify incidents involving both tactics. The RSA Pilot Test DIF classification gave priority to the force tactic when both

Table 7-15. Comparison of prevalence rates from RSA Pilot Test to National Intimate Partner and Sexual Violence Survey (NISVS) rates, by type of victimization, reference period, mode, and classification method for females ages 18-49 in the general population

	NISVS ^{f*}	Past 12 months ^a			
		RSA Pilot Test ^c			
		CATI ^d	SV ^g	DIF ^h	ACASI ^e
Rape (including forced attempts) ^o	2.0 %	~	2.7 %	~	2.7 %
Completed rape ⁱ	1.8 %	2.2 %	2.1 %	2.8 %†	2.4 %
Completed forced penetration ^j	0.8	1.3	1.7 †	1.7 †	1.9 †
Completed alcohol/drug facilitated penetration ^k	1.2	1.1	0.5 †	1.2	0.5 †
Attempted forced penetration ^p	0.6 %	~	.9 %	~	1.2 %†
Completed coerced penetration ^l	0.1 %!	0.3 %	0.1 %!	0.3 %	0.2 %!
Completed unwanted sexual contact ^{m,n}	3.7 %	5.9 %†	4.8 %	7.5 %†	5.5 %†
Completed unwanted sexual non-contact ^m	2.7 %	1.6 %†	0.9 %†	2.1 %	1.1 %†

Table 7-15. Comparison of prevalence rates from RSA Pilot Test to National Intimate Partner and Sexual Violence Survey (NISVS) rates, by type of victimization, reference period, mode, and classification method for females ages 18-49 in the general population (continued)

	NISVS ^{f*}	Lifetime ^b	
		RSA Pilot Test ^c	
		CATI ^d	ACASI ^e
		SV ^g	SV ^g
Rape (including forced attempts) ^o	21.7 %	~	~
Completed rape ⁱ	19.6 %	21.8 %	20.8 %
Completed forced penetration ^j	13.9	17.9 †	17.3 †
Completed alcohol/drug facilitated penetration ^k	10.7	9.8	9.5
Attempted forced penetration ^p	5.5 %	~	~
Completed coerced penetration ^l	1.5 %	5.9 %†	5.1 %†
Completed unwanted sexual contact ^{m,n}	29.8 %	38.6 %	~ %
Completed unwanted sexual non-contact ^m	29.9 %	19.8 %	~ %

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NISVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

~ Not applicable.

^aIncludes incidents that occurred within the past 12 months of the interview.

^bIn the RSA Pilot Test, lifetime victimization was estimated using SV screener questions only.

^cBased on females ages 18-49 in the general population sample of the RSA Pilot Test, 2014-2015 .

^dComputer-assisted telephone interview.

^eAudio computer-assisted self-interview.

^fNational Intimate Partner and Sexual Violence Survey 2010. The NISVS uses a national probability sample.

^gSexual victimization screener. This method uses behaviorally specific questions to classify incidents. All SV estimates are after de-duplication process.

^hDetailed incident form. This method uses multiple questions from the detailed incident form to classify incidents. All DIF estimates count only the most serious behavior and tactic for each incident.

ⁱIncludes penetrative sexual contact using force or while unable to consent where the offender completed the behavior.

^jIncludes penetrative sexual contact using force where the offender completed the behavior.

^kIncludes penetrative sexual contact using while unable to consent where the offender completed the behavior.

^lIncludes penetrative sexual contact using coercion where the offender completed the behavior.

^mLifetime estimates not available for ACASI respondents.

ⁿIncludes unwanted kissing, grabbing, or fondling of breasts, genitals or buttocks using any tactic where the offender completed the behavior.

^oIncludes penetrative sexual contact using force where the offender completed the behavior or while unable to consent where the offender either attempted or completed the behavior.

^pIncludes penetrative sexual contact using force where the offender physically attempted but did not complete the behavior.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015 and Center for Disease Control and Prevention, National Intimate Partner and Sexual Violence Survey (NISVS), 2010.

occurred during the same incident, whereas the NISVS counts the same incident in both rows of the table (see section 8.3 on the overlap between force and inability to consent for the same incidents). A second reason for the difference is that many of the initial responses to the alcohol/drug question on the RSA Pilot Test were eventually classified by the DIF in another category. Both of these reasons may explain why the RSA Pilot Test for the SV measure is closer to the NISVS than the DIF measure. Classification is based on which screening item was used to report the victimization, rather than on more detailed questions about the use of force and the role that alcohol/drugs had on the occurrence of the incident.

It is also possible to compare the two surveys for lifetime estimates. For the RSA Pilot Test, only estimates from the SV screener are available.³⁶ A similar pattern is seen in the lifetime rates for completed rape as discussed above for the 12-month estimates. Overall estimates of lifetime completed rape are not significantly different between NISVS and the RSA Pilot Test. As with the 12-month estimates, the RSA Pilot Test found higher rates for rape by force for both modes. These differences are statistically significant, whereas for the 12-month comparisons for the SV screener only one was significant. Also, just as with the 12-month comparisons, rates of lifetime completed alcohol- or drug-facilitated rape are not significantly different from the NISVS.

The different patterns by physical force may be related to the order of the screening items. The NISVS first asks about alcohol/drug-facilitated penetration and then about physical force. The RSA Pilot Test reversed this order.

Coerced penetration. Both of the surveys ask about penetration due to coercion, which is force using non-physical tactics. The operationalization of the concepts on the two surveys is slightly different. The NISVS definition is –

being pressured in ways that included being worn down by someone who repeatedly asked for sex or showed they were unhappy; feeling pressured by being lied to, being told promises that were untrue, having someone threaten to end a relationship or spread rumors; and sexual pressure due to someone using their influence or authority. (Black et al., 2011, p. 17)

³⁶The RSA Pilot Test respondents were asked to fill out a DIF only for incidents occurring with the last 12 months.

The RSA Pilot Test screening item initially asks, “Has anyone made you have any type of sex by threatening to cause problems for you, such as at your job or school, at home, in your relationships, or in any other way?” (SV6 screener item). To be classified as coercion on the DIF, the respondent had to say “yes” to one of the following items:

- D3B threaten to cut off financial support
- D3C threaten to cause problems at your job, at school, in your relationships, or to cause some other problem.
- D3D promise rewards in your relationship, your job, your grades or something else.

The NISVS definition is a bit broader because it includes verbal pressure. The types of examples that are cited in each survey are also a bit different, although they cover very similar territory, such as threats to relationships and use of authority to either punish or reward the victim.

Rates of completed coerced penetration in the past 12 months are very low for both studies. There are no significant differences between NISVS and the RSA Pilot Test estimates for victimizations in the last 12 months. However, NISVS rates were significantly lower than RSA Pilot Test estimates for lifetime victimizations.

Unwanted Sexual Contact. The NISVS asks about the same set of unwanted non-penetrative sexual contact, as did the RSA Pilot Test. For purposes of comparison, the RSA Pilot Test included non-penetrative sexual contact regardless of tactic. The NISVS has rates of completed unwanted sexual contact at 3.7 percent in the past 12 months, which were significantly lower than three of the four estimates from the RSA Pilot Test, which ranged from 4.8 percent to 7.5 percent. The differences are larger for the lifetime estimates, where the NISVS estimate is about 9 percentage points lower than the SV estimate for the RSA Pilot Test (29.8% vs. 38.6%).

Unwanted Sexual Non-Contact. The two surveys also asked about unwanted sexual non-contact. The NISVS defines these as –

those unwanted experiences that do not involve any touching or penetration, including someone exposing their sexual body parts, flashing, or masturbating in front of the victim, someone making a victim show his or her body parts, someone making a victim look at or participate in sexual photos or movies, or someone harassing the victim in a public place in a way that made the victim feel unsafe.

The RSA Pilot Test definition was similar to that of the NISVS except it did not include the harassment component.

NISVS rates of unwanted sexual non-contact in the past 12 months were significantly *higher* than three out of the four RSA Pilot Test estimates. NISVS estimated the prevalence rate of unwanted sexual non-contact to be 2.7 percent, and the RSA Pilot Test estimates ranged from 0.9 to 2.2 percent. This difference is also reflected in the lifetime measures. The NISVS is about 10 percentage points higher than the RSA Pilot Test (29.9% vs. 19.8%). The inclusion of harassment in the NISVS definition may account for much of this difference. A second reason may be because the NISVS asked about these incidents at the beginning of the sexual violence portion of the survey instrument, whereas the RSA Pilot Test asked these as the last set of items on the SV.

7.2.3 Summary of Results on Prevalence and Incidence

When finalizing the classification algorithm, the narratives were used to assess measurement issues with the detailed incident form. This review indicated there were measurement issues related to identifying attempts and threats. In a significant number of cases, respondents were reporting verbal threats, which did not pose an immediate and credible threat. Many of these may be considered harassment, rather than a threat of rape or sexual assault. The DIF did not include follow-up probes asking about what constituted the actual attempt or threat. More data are presented in Chapter 9 on this topic, as well in the final chapter on recommendations.

The general population estimates of prevalence for ACASI and CATI for the RSA Pilot Test for rape and sexual assault was 5.9 percent and 5.3 percent for women ages 18-49 living in the five targeted CBSAs. The rate of rape was slightly higher than for sexual assault. For example, for ACASI, 3.8 percent reported being raped at least once, while 2.8 percent reported being sexually assaulted at least once. Acts occurring by force were more common than those related to inability to consent. Some of this is an artifact of the methods used to classify the incidents, which labeled the incident as using force when both tactics occurred in the same incident. The two-stage design of the survey allows for modifying this rule if so desired. Analysis discussed in section 8.3 takes advantage of this aspect of the design and provides more detail on the number of incidents where both tactics were involved. However, even accounting for this classification rule, the majority of rapes and sexual assaults measured on this survey involved some type of force.

One of the primary hypotheses related to the RSA Pilot Test was to test for a difference between ACASI and CATI administration of the interview. Based on the research literature, it was expected that the self-administered interview (ACASI) would produce higher rates of rape and sexual assault than the interviewer-administered survey (CATI). Tests of this hypothesis described in this chapter are mixed. For the GP sample, the ACASI sample was consistently higher than the CATI, but the differences were small. None of the differences for estimates of rape and sexual assault were statistically significant. There were significant differences for unwanted sexual contact, as well as for respondents reporting no unwanted behavior on the DIF. For the VO sample, however, there were large mode differences, especially for rape. Some of the differences, even those that were not statistically significant, may be due to greater external telescoping on the ACASI. This is discussed in section 9.1.2. However, this explanation does not account for the large differences for not only rape, but other, less salient events, such as unwanted sexual contact, which are less susceptible to this type of recall error.

One difference between the VO and GP samples was that the VO group reported many more victimizations, even after controlling for age and other characteristics. This suggests that the differences between the ACASI and CATI may be accentuated for those with the most to report. In a recent study comparing interactive voice response and CATI, a similar result was found. A special sample of households that had reported a crime to the police were randomly assigned to each mode. This sample exhibited much higher victimization rates when compared to the general population sample. They also exhibited much larger mode effects than the general population sample (Cantor & Williams, 2013).

Some of this difference may also be related to the non-random nature of the VO sample. These individuals volunteered to participate in the survey by responding to a Craigslist ad. Their motivation was, at least initially, related to the incentive. A personal visit by an interviewer may have instilled more motivation for them to try to recall and report on incidents than those participating by telephone.

Regardless of the explanation, this mixed result suggests that if there is a mode effect, it is not large for a general population survey of women ages 18-49. The privacy and confidentiality protections enforced during the telephone interview are comparable, at least to a large degree, to that afforded by the ACASI. Confidentiality and privacy have several dimensions. One is reluctance to talk about an incident to another person because of embarrassment or shame. Another is reluctance to report an incident for fear that knowledge of the victimization by another person will cause some damage or harm to the respondent (shame; personal consequences; being arrested for

illegal behavior). The above results are consistent with the second explanation having to do with the personal consequences of someone the respondent knows learning about the incident.

This explanation is consistent with several other recent findings. Cantor and Williams (2013) found that a self-administered interview over the telephone using an interactive voice response (IVR) did not yield more estimates of rape and sexual assault than a CATI interview. Laaskonen and Heiskanen (2014) compared self-administered and interviewer administered modes for sexual violence. They did not find a statistically significant difference for the sexual violence question; although the difference was in the expected direction (IVR is higher).

This chapter also compared the RSA Pilot Test estimates to those of both the NCVS and NISVS. With respect to the NCVS, the RSA Pilot Test estimates of incidence were at least 50 times higher than those of the NCVS. The study did not experimentally vary design features, so it is difficult to account for exactly why the estimates are so different. It was possible to simulate two sources of the difference. One was the expanded scope of incidents that are asked about on the RSA Pilot Test relative to the NCVS. A second was the context of the two surveys. The NCVS is presented as a survey about crime, while the RSA Pilot Test was introduced as one on health and safety. Once restricting the RSA Pilot Test to comparable content and to incidents respondents thought were a crime at the time of the incident, the difference between the surveys was significantly reduced, but there still remained a significant difference between the two surveys.

The rates for completed rape on both the NISVS and the RSA Pilot Test were very similar. The rates for the RSA Pilot Test were nominally higher, but only a few were statistically higher. The rates of lifetime victimization were also very close. Although the estimates do not cover the same geographical areas or the same time period, it is interesting that the rates are as close as they are given the differences in the methodologies. The NISVS is oriented toward intimate partner violence, whereas the RSA Pilot Test asks about incidents of different types of unwanted sexual misconduct, without reference to a partner. Both of the surveys use BSQs, but the context, order, and specific language differs. The NISVS includes a number of sections that cover violence and other predatory incidents before asking about rape and sexual assault (e.g., stalking; physical violence; intimate partner violence). In terms of the order of the rape and sexual assault questions, the NISVS starts with non-penetrative acts and moves to alcohol/drug-facilitated penetration and then to penetration by force. The opposite is the case for the RSA Pilot Test. In terms of wording, the RSA Pilot Test uses the phrase “against your will” to describe nonconsent for most of the questions (the exception being the alcohol/drug question), while the NISVS uses “when you didn’t want it to happen.”

The most significant empirical difference between the surveys is the relative occurrence of incidents involving force and drug/alcohol use. The NISVS has more incidents involving alcohol and drugs than due to force. The RSA Pilot Test estimates based on the DIF classify significantly more by force than inability to consent. Part of this is because of the classification methods used by the DIF on the RSA Pilot Test. There are a number of incidents involving both force and inability to consent that are classified as force. However, there are also a number of incidents that are screened as rape when the victim was unable to consent but are reclassified into another category once the DIF items are administered. Section 9.2.1 provides more detail on why these two sets of estimates may differ.

For estimates of coercion and non-penetrative unwanted sexual contact, the NISVS rates were consistently below those for the RSA Pilot Test. The opposite is the case for estimates of unwanted sexual non-contact. We speculate that this was due to two factors. One is the wider definition used on the NISVS, which included verbal harassment. The second is the order of the screening questions, which asked these items first on NISVS and last on the RSA Pilot Test.

8. Characteristics of Victims and Incidents

This chapter describes the characteristics of the victims that were identified on the RSA Pilot Test, the types of incidents that were reported and how these compare to the victims identified on the NCVS.

8.1 Characteristics of Victims

The RSA Pilot Test collected information on the demographic and economic characteristics of respondents, such as age, race/Hispanic origin, education, college enrollment, income, marital status, employment, home ownership, and geographic mobility. The survey also collected information on activities that might be related to risk of victimization, such as going out at night, going out during the day, and the use of public transportation. These characteristics have been hypothesized to be related to victimization. This section describes the relationship between the RSA Pilot Test measures of rape and sexual assault with each of these characteristics.

Prevalence rates were estimated for the above characteristics for rape, sexual assault, and other unwanted sexual contacts for the general population sample. These rates are provided in Appendix F. To summarize the relationships between victimization and each characteristic, logistic regressions were estimated predicting the prevalence of rape and sexual assault with three variables—mode, the characteristic, and the interaction between the two. The results of these regressions are provided in Appendix F. Table 8-1 shows which respondent characteristics are significantly predictive of rape and sexual assault rates, and which interact significantly with mode of interview. In the remainder of this section, the results for the relationships that were statistically significant are described in more detail.

Table 8-1. Significant predictors of rape and sexual assault rates, by characteristic, mode, and interaction

Characteristic	Main effect of characteristic		Main effect of mode		Mode x characteristic interaction	
	Rape	Sexual assault	Rape	Sexual assault	Rape	Sexual assault
Race/Hispanic origin	✓				✓	✓
Age	✓	✓			✓	
Highest grade completed		✓				
Currently in school and sorority	✓			✓		
Total household income during 2013			✓	✓		
Marital or relationship status	✓	✓				
Had a job the week before interview		✓				✓
Own or rent home		✓			✓	✓
Length of time lived at current address						
Average frequency of shopping last 12 months						
Average frequency spent the evening out	✓				✓	
Average frequency of public transportation use			✓			✓

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Race and Hispanic Origin. Respondents were separately asked their race and their Hispanic origin. This was coded into six categories (Hispanic, and non-Hispanic white, non-Hispanic black, Asian, other, and more than one race). Hispanics had significantly lower rates of rape when compared to white respondents. There was also a difference between the modes with respect to Asians. For ACASI, this group had very low rates of victimization, while it was about average for the CATI.

Age. There is a negative relationship between age and both rape and sexual assault. For rape, for example, women ages 18-20 had a rate of 9.5 percent in the ACASI sample compared to 3.9 percent for those ages 25-29 and 2.3 percent for those ages 30-49. A similar decrease in rates is observed for sexual assault across both modes of interviewing.

Highest Grade Completed. Respondents' highest level of education was grouped into five categories: (1) no high school diploma, (2) high school graduate or GED, (3) some college, associate's degree, or vocational, (4) bachelor's degree, and (5) graduate degree. This is related to sexual assault rates for both modes. Those without a high school degree have the lowest rates (e.g., 0.8 for no high school diploma vs. 3.3 for with a high school diploma; ACASI).

Current Enrollment Status. College attendance is thought to affect risk of sexual violence. One measure collected on the survey was whether the respondent was currently enrolled in college. If enrolled, she was also asked if she was in a sorority. Respondents were classified into four categories: (1) not currently in school, (2) currently in college, not in a sorority, (3) currently in college and in a sorority, and (4) currently in school but not college (e.g., high school, vocational school). For rape, those currently in college had significantly higher rates than those not in college. There is no difference by whether they are in a sorority. For assault, being in a sorority does elevate the risk of victimization.

Household Income. Income is thought to be related to the risk of predatory crimes, including rape and sexual assault. The results for household income differed by type of assault and mode of interview. For rape, the relationship was different for the two interview modes. For the ACASI, those in the lowest income group had the highest rates (6.9% for <\$25,000 vs. 3.6% or lower for those with at least \$25,000). There was no discernible pattern across income groups for the respondents who took the survey on CATI. There was not a significant relationship for sexual assault for either mode.

Marital Status or Relationship Status. As with income and age, marital status and involvement with romantic partners are thought to be related to sexual violence. Respondents were asked for their current marital status, if they were cohabitating with someone else, and if they have had a romantic relationship in the last year. A single variable was created that had four categories: (1) currently married, (2) not married and living with a partner, (3) not married, not cohabitating, and had a romantic relationship in the past year, and (4) not married, not cohabitating, and did not have a romantic relationship in the past year. Those that are currently married had significantly lower rates of rape when compared to the other three categories of not-married women. Those who had a romantic relationship in the past year had the highest rates. For sexual assault, the relationships were similar but not identical. Married women had a significantly lower rate of sexual assault than those who had been in a romantic relationship.

Employment. Respondents were asked if they were employed the week before the interview. This was not related to rates of rape. For sexual assault the relationship differed by mode. For the CATI interviews, those that did have a job the previous week were more likely to be victims of sexual assault than those that did not have a job. Employment in the last week was not significantly related to sexual assault for the ACASI mode.

Home Ownership. On the NCVS, home ownership is positively related to several different types of victimization. It is an indicator of both residential stability as well as economic status. For rape, the results are different by mode of interview. For those responding by ACASI, renters are more likely to report a rape than homeowners (5.3% vs. 2.0%). This relationship is not statistically significant for the CATI. For sexual assault, renters have significantly higher rates than owners for both modes.

Frequency Spent the Evening Out. The survey asked respondents how often they spent the evening out. The response categories went from “almost every day” to “never.” This was one of several measures of how often the respondent may be exposed to potential offenders outside the home. This characteristic was significant for rape, but the results were different by mode of interview. For those responding on the CATI, there is a significant negative relationship between frequency of going out and rape victimization. Those going out almost every day have a prevalence rate of 5.0 percent compared to those who go out less than once a month (0.9 percent). For ACASI, this relationship is scattered with no clear pattern. Those who go out less often have the highest rates (5.9%), with the other rates being somewhat lower.

Frequency of Taking Public Transportation. Taking public transportation was a second measure of being exposed to potential offenders outside the home. The relationship differs by mode of interview. For the ACASI, those who use public transportation the least have the lowest rates. This relationship is not significant for those taking the CATI interview.

A number of the above variables are inter-correlated. For example, age is correlated with the frequency of nighttime activities and college enrollment. To account for these correlations, a step-wise logistic regression was estimated that entered each of the variables listed in table 8-1, including the significant interactions, to separately predict rape and sexual assault. The significant predictors are shown in tables 8-2 and 8-3, along with the associated odds ratios. For rape, the significant predictors are age, race/Hispanic origin, marital status/relationship, and income. The direction of the relationships is consistent with what was discussed above, with one exception. For race/Hispanic origin, non-Hispanic whites have the highest rate of sexual assault.

Several of the significant bivariate relationships do not carry over to the summary model, such as college enrollment, home ownership, and the indicators of activities outside the home.

For sexual assault, age, race/Hispanic origin, marital status/relationship, and education remained significant. However, age is not as significant as for rape. While the odds ratios decrease as age increases, the only coefficient that is statistically different from the youngest age group is the oldest age group (40-49). The direction of these effects is the same as the bivariate results with one exception. As with the summary model for rape, non-Hispanic whites have the highest risk of sexual assault once controlling for all of the above characteristics.

Table 8-2. Odds ratios for significant predictors of rape victimization^a

	Odds ratio	95% lower bound	95% upper bound
Intercept			
Age			
18-20 (reference group)			
21-24	0.49	0.25	0.94
25-29	0.40	0.23	0.67
30-39	0.24	0.13	0.46
40-49	0.19	0.10	0.37
Marital status and relationship			
Married (reference group)			
Not married, living with a partner	2.31	1.20	4.46
Not married, not living with partner, rr ^b	5.16	3.03	8.79
Not married, not living with partner, no rr ^a	2.42	1.27	4.63
Race/Hispanic origin			
Non-Hispanic white (reference group)			
Non-Hispanic black	0.54	0.34	0.87
Non-Hispanic Asian	0.40	0.05	3.12
Non-Hispanic other	1.23	0.44	3.43
Non-Hispanic more than one race	0.83	0.25	2.73
Hispanic	0.30	0.19	0.47
Income			
Less than \$25,000 (reference group)			
\$25,000 – \$49,999	0.65	0.42	1.01
\$50,000 – \$75,000	0.69	0.39	1.23
More than \$75,000	0.29	0.16	0.53
Likelihood ratio Chi Square 51406.1 p< .0001			
^a regressions use combined data for ACASI and CATI			
^b rr – romantic relationship in past 12 months			
Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015			

Table 8-3. Odds ratios for significant predictors of sexual assault victimization^a

	Odds ratio	95% lower bound	95% upper bound
Intercept			
Age			
18-20 (reference group)			
21-24	1.05	0.58	1.92
25-29	0.70	0.37	1.35
30-39	0.59	0.30	1.14
40-49	0.44	0.23	0.83
Marital status and relationship			
Married (reference group)			
Not married, living with a partner	1.89	0.96	3.72
Not married, not living with partner, rr ^a	4.25	2.66	6.79
Not married, not living with partner, no rr ^a	1.82	1.01	3.26
Race/Hispanic origin			
Non-Hispanic white (reference group)			
Non-Hispanic black	0.61	0.37	1.00
Non-Hispanic Asian	0.42	0.21	0.87
Non-Hispanic other	0.38	0.11	1.35
Non-Hispanic more than one race	0.58	0.19	1.77
Hispanic	0.45	0.29	0.69
Education			
No high school diploma (reference group)			
High school graduate or GED	2.69	1.06	6.84
Some college, associate's degree, or vocational school	4.24	1.79	10.02
Bachelor's degree	2.70	1.10	6.61
Graduate degree	2.47	0.87	7.00
Likelihood Ratio Chi Square 36187.7 p< .0001			
^a rr – regressions use combined data for ACASI and CATI			
^b rr = romantic relationship in past 12 months			
Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.			

8.2 Circumstances of the Incident

One of the key features of a two-stage design is the collection of incident-level information. The characteristics collected on the RSA Pilot Test include the time and location of the incident, the behaviors and tactics used during the incident, if injuries were sustained, offender characteristics, the role of alcohol and drugs during the incident, any follow-up with mental health professionals, and follow-up with the police, service providers, or friends.

Detailed incident level information was collected on up to three incidents, with the most serious and most recent incidents taking priority. The first DIF completed by a respondent included the full set of questions, whereas subsequent DIFs included a subset of the items. A shortened DIF was used to reduce respondent burden. The questionnaire in Appendix F can be referenced for a comprehensive list of items administered on the short and long versions of the DIF. With the exception of two characteristics, the discussion below provides data taken from those items that were included on both the long version of the DIF.

The content of DIF was identical across the two modes of interviewing; however, all questions on the telephone survey were designed with either a yes/no response or by asking for a response to the number corresponding to the particular response category. This was to make it difficult for anyone who did overhear the interview to understand what was being discussed. In CATI, respondents could always volunteer a “don’t know” or “prefer not to answer” response. In ACASI, respondents could leave a question blank.³⁷

The discussion below concentrates on the GP sample. The results for the VO sample are provided in Appendix F.

8.2.1 Time of Day of Incident

The DIF began by asking respondents about the time of day the incident occurred. Approximately three-quarters of all incidents of sexual victimization occurred at night and only about one-quarter during the daytime (table 8-4). This finding was consistent across sample types and different types of victimization. The distribution of daytime versus nighttime is nearly identical between the modes. In terms of the specific time of day, there is much more missing data for the ACASI. This is primarily due to the presentation of a “don’t know” category as a response option. CATI respondents were not given this option unless they specifically said they refused or did not know.

³⁷If a respondent left a screen blank, she was then asked if she didn’t know the answer, refused to answer, or had unintentionally skipped the item.

Table 8-4. Time of day incident occurred, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Daytime	25.9 %	28.3 %	21.8 %	20.6 %
6 a.m. to noon	4.6	7.9	6.0	9.8
Noon to 6 p.m.	14.5	18.6	10.2	10.3
Do not know time of day	6.8 †	1.7	5.6 †	0.5 !
Nighttime	74.1 %	71.7 %	78.2 %	79.4 %
6 p.m. to midnight	35.3	39.1	35.2	35.8
Midnight to 6 a.m.	20.7 †	30.8	24.6 †	41.7
Do not know time of night	18.1 †	1.8	18.4 †	1.9 !
Number of weighted sample cases	2,639,672	1,896,799	568,178	431,944

Table 8-4 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Daytime	25.2 %	32.1 %	27.4 %	30.0 %
6 a.m. to noon	1.5 !	6.7	4.7	7.6
Noon to 6 p.m.	11.1	21.4	16.7	20.9
Do not know time of day	12.7	4.0	6.0 †	1.4
Nighttime	74.8 %	67.9 %	72.6 %	70.0 %
6 p.m. to midnight	37.6	44.6	34.8	38.6
Midnight to 6 a.m.	27.8	21.9	18.0 †	29.4
Do not know time of night	9.4 †	1.4 !	19.8 †	2.0
Number of weighted sample cases	362,621	366,564	1,708,873	1,098,291

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.2 Location of Incident

Incidents took place in a variety of locations, with some differences depending on the type of victimization (table 8-5). About half of the rapes occurred “while hanging out at someone’s home.” Combined with those that occurred at the respondent’s home, around 70 percent of the incidents occurred in someone’s private residence—the victim’s or someone else’s. This differs somewhat from sexual assault and other unwanted contacts in which about one-third occurred at someone else’s private residence and about one-quarter occurred in a public place like a restaurant, bar, or movie.

There were no mode differences.

Table 8-5. Activity at time of incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

Activity at time of incident	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
In transit	11.8 %	13.7 %	7.1 %	8.8 %
Hanging out at someone’s home	27.4 %	26.3 %	37.5 %	30.6 %
At a restaurant, bar, or movie	24.3 %	25.0 %	9.2 %	10.4 %
Working or at school	11.4 %	11.0 %	10.2 %!	6.2 %
Hanging out at your home	20.9 %	19.3 %	33.3 %	37.0 %
At an outdoor public space	0.9 %!	1.1 %	0.7 %!	1.0 %!
Sleeping (location not specified)	0.2 %!	0.8 %!	0.5 %!	2.4 %!
In a vehicle (not in transit)	-- %	0.1 %!	-- %	--
Other	3.1 %	2.6 %	1.5 %!	3.7 %!
Number of weighted sample cases	2,668,490	1,948,925	572,188	477,462

Table 8-5. Activity at time of incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Activity at time of incident				
In transit	7.1 %	11.9 %	14.4 %	16.5 %
Hanging out at someone's home	25.1 %	26.8 %	24.6 %	24.3 %
At a restaurant, bar, or movie	39.5 %	36.1 %	26.1 %	27.5 %
Working or at school	12.0 %	9.1 %	11.6 %	13.8 %
Hanging out at your home	11.4 %	10.6 %	18.8 %	14.6 %
At an outdoor public space	-- %	3.2 %!	1.1 %!	0.4 %!
Sleeping (location not specified)	-- %	--	0.1 %!	0.4 %!
In a vehicle (not in transit)	-- %	--	-- %	0.2 %!
Other	4.9 %	2.3 %!	3.3 %	2.2 %
Number of weighted sample cases	372,166	375,381	1,724,135	1,096,081

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

-- Less than 0.05%.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.3 Use of Verbal Pressure and Coercive Tactics In Incident

Respondents were asked whether the offender engaged in various tactics that did not involve physical force at any point leading up to the incident, including verbally pressuring or continually arguing with the woman; and they were asked about tactics considered to be coercion, including threatening to cut off financial support; threatening to cause problems at her job, at, school, in her relationships or to cause some other problem; and promising rewards in the relationship, job, grades, or something else.

Among incidents reported in the general population sample, more than 4 in 10 unwanted incidents involved the offender using verbal pressure or continual arguments to engage in the behavior (table 8-6). Fewer than 2 in 10 incidents involved one of the tactics considered to be

coercion. Verbal pressure and all forms of coercion were more common in incidents involving rape than those involving sexual assault or other types of incidents.

Rape and sexual assault are defined as involving some type of physical force or inability to consent. The non-physical tactics listed here, therefore, are in addition to these. Among rape incidents, the most common non-physical tactic was verbal pressure. Fully three-fourths of ACASI respondents and nearly 6 in 10 CATI respondents reported this tactic. The difference between modes was statistically significant. Among rape incidents, ACASI respondents were more than twice as likely to report promise of rewards in exchange for the sexual activity than were CATI respondents.

For sexual assault and other incidents, fewer non-physical tactics were reported. The most common was verbal pressure, which occurred around 45 percent of the time for sexual assault and 35 percent for other incidents. Note that not all respondents who reported an unwanted behavior marked any of the items in this series.

Table 8-6. Use of verbal pressure and coercive tactics, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Type of tactic used in incident				
Verbally pressure you or continually argue with you	46.7 %	43.5 %	74.4 %†	58.7 %
Threaten to cut off financial support	3.8 %	4.9 %	12.2 %	9.6 %
Threaten to cause problems at job, school, relationships	15.0 %	14.0 %	37.7 %	28.0 %
Promise rewards in relationship, job, grades, something else	15.0 %†	7.2 %	26.2 %†	10.7 %
Number of weighted sample cases	2,285,296	1,822,356	561,230	477,462

Table 8-6 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Type of tactic used in incident				
Verbally pressure you or continually argue with you	42.4 %	45.3 %	36.4 %	35.2 %
Threaten to cut off financial support	1.9 %!	4.0 %!	0.9 %!	2.9 %
Threaten to cause problems at job, school, relationships	10.7 %	7.2 %	6.9 %	9.7 %
Promise rewards in relationship, job, grades, something else	8.7 %	2.9 %!	12.0 %	7.2 %
Number of weighted sample cases	370,048	375,381	1,354,018	969,512

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.4 Use of Force

Respondents who reported any type of unwanted behavior were also asked whether the offender used any type of force during the incident, including holding or pinning you so you had difficulty moving; using a weapon or threatening to use a weapon; physically attacking you or threatening to attack you but not with a weapon; physically attacking or threatening to attack someone else; or using any other type of physical force.

In more than one-third of incidents, respondents reported that the person held or pinned them so they had difficulty moving (table 8-7). This was the most common type of force for rape. The use of this tactic was as high as 84.8 percent in ACASI and 72.6 percent in CATI. In rape incidents, 3 in 10 also reported that the person physically attacked them or threatened to physically attack them. In cases of sexual assault, more than half of incidents involved holding or pinning, and at least one-fourth of incidents involved grabbing, pushing, or pulling the woman. There were no mode differences among the general population.

Several types of force were reported in the “other specify” category that were classified as force but were not explicitly covered in the response categories. One was when the respondent was groped or grabbed without warning. The second were instances where the respondent was trapped in a space, like a car or room and could not get out. Some consideration should be given to including these categories as types of force in the future.

Table 8-7. Force tactics, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^e	CATI ^f *	ACASI ^e	CATI ^f *
Type of force used in incident				
Hold you or pin you so you had difficulty moving	35.9 %	35.7 %	84.8 %	72.6 %
Use a weapon, or threaten to use a weapon	4.3 %	3.6 %	14.8 %	10.8 %
Physically attack you or threaten to attack you, but not with a weapon	12.3 %	13.7 %	31.0 %	32.2 %
Physically attack, or threaten to attack, someone else	4.8 %	5.5 %	13.4 %	12.0 %
Grabbed/pushed/pulled	5.8 %	9.6 %	0.3 %!	1.6 %!
Other type of force	1.5 %	0.8 %!	0.9 %!	1.7 %!
Number of weighted sample cases	2,330,449	1,822,356	574,473	477,462

Table 8-7 (continued)

	Sexual assault ^b		Other incidents ^{c,d}	
	ACASI ^e	CATI ^f *	ACASI ^e	CATI ^f *
Type of force used in incident				
Hold you or pin you so you had difficulty moving	50.4 %	53.8 %	11.2 %	10.7 %
Use a weapon, or threaten to use a weapon	2.0 %!	4.0 %	0.6 %!	--
Physically attack you or threaten to attack you, but not with a weapon	14.5 %	16.0 %	3.8 %	3.8 %
Physically attack, or threaten to attack, someone else	5.6 %	7.0 %	1.0 %!	1.8 %
Grabbed/pushed/pulled	24.8 %	31.5 %	2.9 %	5.0 %
Other type of force	5.0 %	0.5 %!	0.7 %!	0.4 %!
Number of weighted sample cases	381,836	375,381	1,374,140	969,512

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dSome incidents classified as other unwanted sexual contact do involve force, because incidents of forced contact were not classified as sexual assault if the offender stopped when they said no. See chapter 7.1 for more information.

^eAudio computer-assisted self-interview.

^fComputer-assisted telephone interview.

-- Less than 0.05%

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.5 Injury From Incident

A physical injury was sustained in roughly 8 percent of the incidents; however, the rate of injury was much higher in rape incidents (32.7% ACASI and 24.7% CATI) (table 8-8). Those who experienced injuries were asked to categorize the types of injury. There were no significant differences by mode in the type of injury among rape incidents, though across all types of incidents, there were significantly more injuries from sexual intercourse reported in ACASI than in CATI (table 8-9). In more than half of rape incidents where an injury occurred, the victim received medical care as a result of the incident (54.7% ACASI and 68.3% CATI) (table 8-10).

Table 8-8. Physical injuries, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Any physical injuries as a result of the incident?				
Yes	8.1 %	8.4 %	32.7 %	24.7 %
Number of weighted sample cases	2,731,734	1,962,587	568,274	468,856

Table 8-8 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Any physical injuries as a result of the incident?				
Yes	6.3 %	9.8 %	0.6 %!	1.1 %!
Number of weighted sample cases	381,836	375,381	1,781,623	1,118,350

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-9. Type of physical injuries, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Type of physical injuries				
Injury from sexual intercourse	45.2 %†	21.5 %	50.4 %	28.1 %
Gunshot or stab wounds	6.7 %!*	1.9 %!	5.7 %!	2.8 %!
Broken bones or teeth knocked out	9.4 %!	6.0 %!	8.8 %!	8.6 %!
Bruises, black eye, cuts, scratches, swelling, chipped teeth	80.4 %	88.2 %	79.6 %	90.2 %
Internal injuries	17.6 %	--	16.1 %	--
Knocked unconscious	14.4 %	2.9 %!	15.2 %	4.2 %!
Other injuries	4.0 %!	6.1 %!	2.4 %!	4.8 %!
Number of weighted sample cases	215,255	165,603	185,571	115,939

Table 8-9 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Type of physical injuries				
a. Injury from sexual intercourse	s	7.9 %!	s	s
b. Gunshot or stab wounds	s	--	s	s
c. Broken bones or teeth knocked out	s	--	s	s
d. Bruises, black eye, cuts, scratches, swelling, chipped teeth	s	78.0 %	s	s
e. Internal injuries	s	--	s	s
f. Knocked unconscious	s	--	s	s
g. Other injuries	s	12.4 %!	s	s
Number of weighted sample cases	23,878	36,953	5,806	12,711

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

s Data suppressed for disclosure reasons.

-- Less than 0.05%

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-10. Medical care among those injured, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Any medical care received as a result of the incident?				
Yes	56.4 %	68.0 %	54.7 %	68.3 %
Number of weighted sample cases	220,516	165,603	185,571	115,939

Table 8-10 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Any medical care received as a result of the incident?				
Yes	s	55.8 %	s	s
Number of weighted sample cases	23,878	36,953	11,067	12,711

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

s Data suppressed for disclosure reasons.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.6 Offender Characteristics

Incidents described by women were overwhelmingly committed by a single, male offender, and in most cases, the offender was known to the respondent (either well-known or a casual acquaintance).

More than 9 out of 10 incidents involved a single offender (table 8-11). There was one statistically significant difference by mode. For rape, a significantly lower percentage of incidents involved a single offender for ACASI (85.2%) compared to CATI (96.0%).

Table 8-11. Number of offenders, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Number of offenders				
One	91.5 %†	96.0 %	85.2 %†	96.0 %
More than one	8.5 %†	4.0 %	14.8 %†	4.0 %!
Number of weighted sample cases	2,702,849	1,955,764	572,981	468,856

Table 8-11 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Number of offenders				
One	93 %	95 %	93.3 %	96.3 %
More than one	7.0 %!	5.0 %	6.7 %	3.7 %
Number of weighted sample cases	381,836	375,381	1,748,031	1,111,527

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

More than 95 percent of incidents across both modes were committed by a male offender (table 8-12). There were no significant differences by mode.

Table 8-12. Sex of offender, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Sex of offender				
Male	95.9 %	96.1 %	97.1 %	95.8 %
Female	4.1 %	3.9 %	2.9 %! ^b	4.2 %!
Number of weighted sample cases	2,472,308	1,874,690	482,989	450,195

Table 8-12 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Sex of offender				
Male	97.5 %	98.3 %	95.2 %	95.5 %
Female	2.5 %! ^b	1.7 %! ^b	4.8 %	4.5 %
Number of weighted sample cases	355,002	356,616	1,634,316	1,067,878

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

To better understand the respondent's relationship to the offender, a typology of the relationship was created that ranged from a complete stranger to a spouse or ex-spouse. There were no significant differences in this typology by mode of interviewing, but as one would expect, there are differences across the type of victimization (table 8-13). In incidents involving rape, fewer than one in five were committed by a complete stranger or someone known to the respondent by sight only (15.6% ACASI, 19.9% CATI). In contrast, many more sexual assault incidents were committed by a stranger or person known by sight only to the respondent (43.8% ACASI, 50.6% CATI).

Approximately 25 to 33 percent of rape incidents were committed by an intimate partner (as defined by NCVS as a spouse/ex-spouse or boyfriend/girlfriend/ex-boyfriend/ex-girlfriend) (25.8% ACASI, 32.9% CATI). Again, in contrast, far fewer sexual assault incidents involved an intimate partner (8.8% ACASI, 13.1% CATI).

Finally, a large proportion of rape incidents involved a person that was known to the respondent, but not an intimate partner. Roughly one-third of the incidents involved a friend or ex-friend, 10 to 18 percent a casual acquaintance, and a small proportion were some other well-known person to the respondent, including relatives. These proportions were roughly comparable, although slightly lower, for sexual assault incidents.

Table 8-13. The offender's relationship to the victim by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Offender relationship				
Stranger	34.5 %	35.5 %	9.4 %	14.4 %
By sight only	6.5 %	5.6 %	6.2 %	5.5 %!
Spouse, ex-spouse	3.7 %	4.8 %	7.4 %	16.9 %
Boy/girlfriend or ex-boy/girlfriend	10.3 %	12.9 %	18.4 %	16.0 %
Friend or ex-friend	27.0 %	24.8 %	33.0 %	32.1 %
Other well-known person (incl. relatives)	9.3 %	7.2 %	8.3 %	5.0 %
Casual acquaintance	8.7 %	9.2 %	17.5 %	10.1 %
Number of weighted sample cases	2,449,566	1,868,139	472,906	450,195

Table 8-13. The offender's relationship to the victim by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Offender relationship				
Stranger ^f	39.3 %	41.9 %	40.7 %	42.4 %
By sight only ^g	4.5 %! [!]	8.7 %	7.0 %	4.5 %
Spouse, ex-spouse ^h	3.3 %! [!]	1.3 %! [!]	2.7 %	0.8 %!
Boy/girlfriend or ex-boy/girlfriend ⁱ	5.5 %! [!]	11.8 %	9.0 %	12.0 %
Friend or ex-friend ^j	24.7 %	22.8 %	25.8 %	22.4 %
Other well-known person (incl. relatives) ^k	12.4 %	5.8 %	8.9 %	8.6 %
Casual acquaintance ^l	10.3 %	7.7 %	5.9 %	9.3 %
Number of weighted sample cases	349,741	353,031	1,626,919	1,064,914

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fIncludes those who indicated that the person was "a stranger you had never seen before."

^gIncludes those who indicated that the person was not a stranger but was known "by sight only."

^hIncludes those who indicated that the person was "a spouse or ex-spouse." In RSA Pilot Test, also includes those who were father of the respondent's child.

ⁱIncludes those who indicated that the person was a boyfriend, girlfriend, ex-boyfriend, or ex-girlfriend.

^jIncludes those who indicated that the person was a friend or ex-friend.

^kIncludes those who indicated that the person was well-known, which includes all relatives other than a spouse.

^lIncludes those who indicated that the person was "a casual acquaintance."

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.7 Offender Use of Alcohol and/or Drugs During the Incident

The questionnaire included a question on whether the offender was using alcohol, drugs, or both. In roughly 7 percent of ACASI incidents and 16 percent of CATI incidents, the respondent said they did not know if the offender was using substances. Once these responses are excluded, roughly two-thirds of all incidents involved substance use by the offender. Most of this was alcohol use (41.8% for ACASI; 44.9% for CATI), followed by use of both drugs and alcohol (3.5% ACASI; 15.6% CATI), with less than 5 percent using just drugs. Substance use varied by type of

victimization (table 8-14). About one-third of the rape incidents (using ACASI estimate) occurred where the offender was only using alcohol, whereas about twice percentage (59.2%) of the sexual assault incidents only involved alcohol. Few mode effects were detected.

Table 8-14. Offender use of alcohol or drugs leading up to incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Offender's use of alcohol or drugs in the hours leading up to incident				
Alcohol	41.8 %	44.9 %	31.5 %	40.3 %
Drugs	3.5 %	3.9 %	6.2 %	8.8 %
Both alcohol and drugs	12.3 %	15.6 %	27.3 %	18.0 %
Neither alcohol nor drugs	42.5 %	35.6 %	35.0 %	33.0 %
Number of weighted sample cases	2,543,287	1,656,115	564,534	438,033

Table 8-14 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Offender's use of alcohol or drugs in the hours leading up to incident				
Alcohol	59.2 %	43.7 %	41.5 %	47.5 %
Drugs	4.4 %!	2.3 %!	2.3 %	2.1 %!
Both alcohol and drugs	11.3 %	20.4 %	7.3 %†	12.8 %
Neither alcohol nor drugs	25.1 %	33.6 %	48.8 %†	37.6 %
Number of weighted sample cases	352,989	309,888	1,625,764	908,194

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.8 Respondent Use of Alcohol and/or Drugs During the Incident

Respondents were classified as unable to consent if they indicated either they were passed out for at least part of the incident (G10) or responding directly they were unable to consent (G12a). To further explore the role alcohol or drugs had on the incident the questionnaire included different measures of intoxication and the victim's ability to consent. In addition to asking whether the respondent was using drugs or alcohol (G5), questions were asked about the respondent's physical condition and any outward signs of intoxication, including whether the person was –

- less able to physically resist (G12b)
- too drunk or high to walk by herself (G13)
- able to communicate clearly with others (G15).

Respondents who said they were using alcohol or drugs were also asked about their ability to consent, including whether the alcohol or drugs –

- led to making decisions that she would not have made otherwise (G12c).

Several measures of the role the perpetrator played in the incident were also asked:

- Were alcohol/drugs given without her knowledge? (G6)
- Was the person trying to get her drunk/high to sexually take advantage of her (G7)
- Was she given alcohol/drugs after she was clearly very drunk/high (G9)

One other, less direct, indicator of alcohol or drug use were collected, asking whether she was able to remember the event (G11). This served as an indicator that the respondent may have been blacked out during the incident.

These different indicators were combined to form four levels of drug/alcohol involvement:

- no substance use
- substance use, but no indications of intoxication or being high
- intoxicated, but the victim was able to consent. Indicated either a physical or cognitive impairment but did not indicate that they were passed out or unable to give consent.

- unable to consent. This was defined as those who were passed out for at least parts of the incident or were unable to give consent due to the use of drugs or alcohol. This was used to define a rape or sexual assault.

Table 8-14a shows the questionnaire items and associated responses that were used for each of the above groups. Categories are listed in order of increasing priority. Incidents where respondents did not give any of the responses listed in the table below were categorized as involving no substance use.

Table 8-14a. Questionnaire items used to determine respondent level of substance use

Questionnaire items by level of substance use		Response
Substance use without intoxication		
G5	Had you been using alcohol or drugs in the hours leading up to the incident?	"Alcohol", "Drugs", or "Both alcohol and drugs"
Intoxication and victim able to consent		
G9	Did (the person/any of the people) who did this to you keep giving you (drinks/drugs/drinks and drugs) after you were clearly very (drunk/high/drunk or high)?	"Yes"
G11	Sometimes using (alcohol/drugs/alcohol and drugs) can make it difficult to remember what might have happened. Which of the following best describes how (alcohol/drugs/alcohol and drugs) affected how much you are able to remember about the incident?	"Not able to remember any part"
G12b	Did the (alcohol/drugs/alcohol or drugs) make you less able to physically resist?	"Yes"
G12c	Did the (alcohol/drugs/alcohol or drugs) lead you to make decisions that you would not have made otherwise?	"Yes"
G13	Were you too (drunk/high/drunk or high) to walk by yourself?	"Yes"
G15	Were you able to communicate clearly to others?	"No"
Unable to consent		
G10	Were you passed out for all or parts of this incident, or are you not sure? By passed out, it means that you were unconscious or asleep because of the (alcohol/drugs/alcohol or drugs).	"Yes, for all" or "Yes, for part"
G12	Did the (alcohol/drugs/alcohol or drugs) make you unable to give consent?	"Yes"

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Approximately one-third of unwanted incidents involved the victim using alcohol or drugs in the hours leading up to the incident (33.2% ACASI and 37.7% CATI) (table 8-15). This includes incidents where the respondent had consumed a substance but was not intoxicated (13.5% ACASI and 17.4% CATI), incidents where the respondent was intoxicated but able to give consent (9.3%

ACASI and 10.6% CATI), and incidents where the respondent was unable to consent (10.4% ACASI and 9.7% CATI). There were no significant differences in levels of substance use by mode of interview.

A significant percentage of the incidents involve victims either unable to consent as defined by the survey or intoxicated but still able to consent. In 22.0 percent of ACASI rape incidents, and 30.2 percent of CATI rape incidents, respondents were unable to consent, while roughly 10 to 12 percent happened while the respondent was intoxicated but still able to consent. These percentages are similar for sexual assault, the exception being that CATI respondents were not as likely to be unable to consent.

Table 8-15. Extent of substance use, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent level of substance use				
No substance use	66.8 %	62.4 %	61.0 %	55.6 %
Substance use, but not intoxicated	13.5 %	17.4 %	4.3 %	s
Intoxicated, but able to consent	9.3 %	10.6 %	12.8 %	9.6 %
Unable to consent	10.4 %	9.7 %	22.0 %	30.2 %
Number of weighted sample cases	2,750,340	1,977,695	574,473	477,462
Respondent level of substance use				
No substance use ^f	56.8 %	60.8 %	70.8 %	65.9 %
Substance use, but not intoxicated ^g	10.9 %	18.9 %	17.0 %	22.3 %
Intoxicated, but able to consent ^h	8.0 %	8.3 %	8.4 %	11.7 %
Unable to consent ⁱ	24.4 %	12.1 %	3.8 %	s
Number of weighted sample cases	381,836	375,381	1,794,031	1,124,851

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

s Data suppressed for disclosure reasons.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fIncludes respondents who did not use any alcohol or drugs in the hours leading up to the incident.

^gIncludes incidents where respondents used alcohol or drugs in the hours leading up to the incident, but were not intoxicated.

^hIncludes incidents where respondents were still able to consent during the incident, but were given drinks or drugs after they were clearly drunk or high; were unable to remember any parts of the incident; or, because of the drugs or alcohol were less able to physically resist, made decisions they would not have made otherwise, were unable to walk by themselves, or were unable to communicate clearly.

ⁱIncludes incidents where respondent was passed out for all or parts of the incident or said that alcohol or drugs made them unable to consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

A significant percentage of the incidents that were classified as either rape or sexual assault involving force also involved inability to consent (Table 8-15a). Of all incidents classified as completed rape by force, 18.1 and 19.6 percent on ACASI and CATI, respectively, occurred when the victim was unable to consent. The classification scheme gave priority to the reports of force when generating the prevalence estimates. Approximately the same percentage involved victims who were intoxicated but still able to give consent.

Table 8-15a. Extent of substance use involved in incident by type of incident, level of intoxication and mode for females age 18-49 in the general population, 2014-2015

	Level of substance use			
	None ^a		Substance use without intoxication ^b	
	ACASI ^e	CATI ^f	ACASI ^e	CATI ^f
Rape ^g	61.0 %	55.6 %	4.3 %!	4.6 %!
Completed ^h	49.2 %	51.8 %	4.0 %!	5.9 %!
Forced ⁱ	58.6	63.3	4.7 !	7.2 !
Unable to consent ^j	--	--	--	--
Sexual assault ^k	56.8 %	60.8 %	10.9 %	18.9 %
Completed ^d	54.5 %	55.2 %	11.6 %	21.9 %
Forced ⁱ	61.5	58.4	13.1	23.2
Unable to consent ^j	s	s	s	s
Other unwanted sexual contact ^l	72.7 %	64.9 %	18.7 %	24.0 %
No unwanted behavior	61.0 %	74.9 %	15.9 %	11.5 %
Rape ^g	12.8 %	9.6 %	22.0 %	30.2 %
Completed ^h	15.6 %	8.1 %	31.2 %	34.2 %
Forced ⁱ	18.6	9.9	18.1	19.6
Unable to consent ^j	--	--	100.0	100.0
Sexual assault ^k	8.0 %	8.3 %	24.4 %	12.1 %
Completed ^d	7.8 %	9.2 %	26.1	13.6 %
Forced ⁱ	8.8	9.8	16.6	8.6
Unable to consent ^j	s	s	s	s
Other unwanted sexual contact ^l	8.6 %	11.1 %	--	--
No unwanted behavior	7.7 %	13.6 %!	15.4 %	--

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

s Data suppressed for disclosure reasons.

^aIncludes incidents where respondent had not been using drugs or alcohol in hours leading up to the incident.

^bIncludes incidents where respondent had been using drugs or alcohol in hours leading up to the incident, but were not visibly intoxicated or unable to consent.

^cIncludes incidents where respondent was visibly intoxicated, but not passed out or unable to consent. This includes incidents where the alcohol/drugs made respondent: unable to remember any of the incident, less able to physically resist, make decisions they would not have otherwise, unable to walk by themselves, unable to communicate clearly, and incidents where respondents were continually given alcohol/drugs after they were clearly drunk/high.

^dIncludes incidents where respondent was passed out for all or parts of the incident or said that alcohol or drugs made them unable to consent.

^eAudio computer-assisted self-interview.

^fComputer-assisted telephone interview.

^gIncludes penetrative sexual contact using force or while unable to consent.

^hIncludes incidents where the offender completed the behavior.

ⁱIncludes holding or pinning, using a weapon or threatening to use a weapon, other physical attacks or threats of physical attacks on respondent or someone else.

^jIncludes incidents where respondents were passed out for all or parts of the incident or were unable to consent due to alcohol or drugs.

^kIncludes non-penetrative sexual contact using force or while unable to consent.

^lIncludes unwanted penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Among the victims of rape and sexual assault who were classified as being unable to consent, about half were passed out for all or part of the incident. The other half said they were unable to consent (table 8-15b). These percentages were the same rape and sexual assault. For sexual assault, there were significantly more ACASI respondents who said they were unable to consent when compared to CATI. As noted above, there were more incidents on the ACASI, in general, that were classified as unable to consent for sexual assault when compared to CATI.

Table 8-15b. Incidents of rape and sexual assault when the victim was unable to consent by classification and mode for females age 18-49 in the general population

	Rape and sexual assault ^a		Rape ^b		Sexual assault ^c	
	ACASI ^d	CATI ^e	ACASI ^d	CATI ^e	ACASI ^d	CATI ^e
Did not involve inability to consent	77.1 %	77.8 %	78.0 %	69.8 %	75.6 %	87.9 %
Involved inability to consent						
Conscious but unable to consent ^f	11.7 %	10.6 %	8.1 %	13.9 %	17.2 %†	6.5 %
Passed out for part of incident	9.8	9.2	11.5	13.8	7.2	3.5
Passed out for entire incident	1.4 !	2.3 !	2.3 !	2.5 !	--	2.1 !
Number of weighted incidents	956,310	852,843	574,473	477,462	381,836	375,381

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

--Less than 0.05%.

^aIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^bIncludes penetrative sexual contact using force or while unable to consent.

^cIncludes non-penetrative sexual contact using force or while unable to consent.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fIncludes incidents where respondents were not passed out for any part of the incident, but said that alcohol or drugs made them unable to give consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

One concern related to measuring alcohol and drug facilitated rape and sexual assault is defining when the victim meets the standard of not being able to give consent. Incidents involving inability to consent are sometimes difficult to prove in a court proceeding and pose unique problems with respect to measuring on a survey. Among those who were considered unable to consent, 91.3 percent reported at least one of the signs of intoxication above (table 8-15c). A large portion (about 80 percent) said they were less able to physical resist or that the alcohol/drugs led her to decisions she would not otherwise make. Almost 70 percent were either unable to walk by herself or unable to communicate clearly. About 42 percent said the other person was giving her drinks after she was clearly drunk. The results were very similar when just looking at these indicators for those who said they were unable to consent. About 90 percent reported at least one sign of intoxication, with the most common indicators being less able to physically resist and making decisions she would not otherwise make. About 60 percent showed outward signs of being intoxicated. These percentages are slightly higher for those who were passed out for some part of the incident.

One of the more common signs reported is the respondent's decisionmaking ability (82.6%). This is a more subjective standard than many of the other signs of intoxication noted in the table. However, even when taking this question out of the calculation, 89 percent of the respondents reported at least one of the other signs (data not shown).

Table 8-15c. Responses to questions on intoxication for incidents when victim was unable to consent for females age 18-49 in the general population, 2014-2015

	All incidents involving inability to consent	Type of inability to consent		
		Conscious but unable to consent ^a	Passed out for part of incident	Passed out for all of incident
Any intoxication item below	91.3 %	90.5 %	96.6 %	s
Person kept giving you drinks after you were clearly drunk	42.5 %	46.5 %	34.0 %	s
Unable to remember any part of the incident	2.5 !	--	5.5	~
Alcohol or drugs made you less able to physically resist	80.9	77.2	85.6	~
Alcohol or drugs led you to make decisions you would not have otherwise	82.6	82.6	82.7	~
Alcohol or drugs made you either unable to walk by yourself, or unable to communicate clearly	66.9	58.4	77.9	~
Unable to walk by yourself	37.5 %	19.3 %	60.7 %	~
Unable to communicate clearly	56.8	50.3	65.3	~
Number of weighted sample cases	477,850	243,740	200,971	33,139

Note: Estimates are based on weighted data. See Appendix A for standard errors.

s Data suppressed for disclosure reasons.

! Interpret with caution. Coefficient of variation is greater than 50%.

~ Not applicable.

^aIncludes incidents where respondents were not passed out for any part of the incident, but said that alcohol or drugs made them unable to give consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.9 Efforts to Try to Stop the Incident

Respondents who experienced any type of force or coercion during the incident, or were unable to consent at the time of the incident, were asked what actions they took during the incident to try to stop it from happening. At this point in the survey, respondents were reminded that there are no right or wrong ways to react in these situations and that these questions would be used to better understand what women in similar situations have done. Respondents were asked to say “yes” or “no” to each action. Respondents reporting on the first incident were asked the full set of reasons, whereas those responding about the second or third incident were asked an abbreviated set of reasons, with an opportunity to list out other reasons not mentioned. These verbatim reasons were then coded to the full set of reasons.

Many women indicated they did multiple things in reaction to the unwanted conduct. The most common responses were to physically resist and/or to say “no,” “stop” or that she did not want the act to happen (table 8-16). The next most common response was to physically resist or try to physically resist the person. For victims of rape, for example, more than eight said they did this. Rape incidents were different from sexual assault in several ways. Rape victims were more likely to try to persuade the offender not to do it, as well as escape or get away. They were less likely (or able) to leave or stop the situation before it actually occurred.

Table 8-16. Respondent actions to try to stop incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent actions to try to stop incident				
Physically resist or try to physically resist the person	74.2 %	79.9 %	84.1 %	79.3 %
Say "no," "stop" or that you didn't want the act to happen	84.4 %	80.0 %	91.6 %	79.6 %
Leave or stop the situation before the act occurred	47.4 %	45.1 %	34.3 %	32.0 %
Attack or threaten the person	15.2 %	23.4 %	18.1 %	19.4 %
Try to persuade, plead or argue with the person	43.6 %	41.3 %	68.4 %	60.8 %
Try to escape or get away	58.2 %	62.7 %	64.4 %	54.5 %
Try to get help	25.9 %	26.0 %	30.6 %	25.9 %
Do something else	4.6 %	6.3 %	5.6 %	5.1 %
Number of weighted sample cases	1,883,234	1,487,428	564,792	464,020

Table 8-16. Respondent actions to try to stop incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Respondent actions to try to stop incident				
Physically resist or try to physically resist the person	75.8 %	85.6 %	68.1 %	77.9 %
Say "no," "stop" or that you didn't want the act to happen	76.3 %	66.2 %	82.9 %	86.2 %
Leave or stop the situation before the act occurred	52.5 %	55.3 %	54.9 %	49.5 %
Attack or threaten the person	19.6 %	25.1 %	11.7 %†	25.6 %
Try to persuade, plead or argue with the person	43.7 %	40.8 %	24.7 %	26.2 %
Try to escape or get away	57.8 %	75.5 %	53.5 %	62.2 %
Try to get help	22.5 %	33.5 %	23.2 %	22.2 %
Do something else	7.0 %!	9.4 %	3.4 %	5.8 %
Number of weighted sample cases	307,378	308,985	1,011,064	714,424

Note: Respondents to the first incident were asked the full set of reasons. For the second and third incidents, respondents were asked the first two reasons and then an open-ended item. Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.10 Did the Offender Immediately Stop When You Said You Didn't Want it to Happen?

Respondents who told the offender “no” or “stop” or that they did not want the incident to happen were asked if the offender stopped immediately when they said this. Among the general population sample, in incidents of rape, only roughly 15 percent of offenders stopped immediately after the respondent told the offender to stop (table 8-17). Less than 10 percent stopped immediately in incidents of sexual assault. It was primarily the “other incidents” that the offender was most likely to stop immediately when asked. This pattern is partly the result of using question G17 (did the offender stop) to classify the incident as a sexual assault (see chapter 5). In particular, for non-penetrative forced sexual contact or attempted forced penetration, if the offender stopped immediately after being asked to do so by the respondent, the incident was classified as unwanted.

Table 8-17. Offender immediately stopped when asked, by type of victimization and mode of interview for females ages 18–49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Offender immediately stopped when asked?				
Yes	50.6 %	42.8 %	14.3 %	16.0 %
Number of weighted sample cases	1,744,888	1,319,259	506,619	358,740

Table 8-17 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Offender immediately stopped when asked?				
Yes	7.6 %	1.4 %!	78.0 %†	67.2 %
Number of weighted sample cases	220,896	209,502	1,017,373	751,018

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

8.2.11 Level of Distress

Victims were asked if they were distressed in the days following the incident and at the time of the interview. In the days following the incident, approximately 4 in 10 incidents led to respondents saying that they were moderately or severely distressed in the days following the sexual victimization (table 8-18). There was a very strong relationship to the type of incident. For victims of rape, roughly two-thirds of respondents were at least moderately distressed. This percentage goes down to roughly 4 in 10 incidents for sexual assault and to fewer than 3 in 10 of the other incidents. There were no significant differences by mode.

Table 8-18. Level of distress in days following incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Level of distress in days following incident				
Not at all distressed	23.0 %	28.8 %	4.9 %	15.5 %
Mildly distressed	33.9	34.8	26.5	21.3
Moderately distressed	25.0	19.9	29.1	28.4
Severely distressed	18.2	16.6	39.4	34.8
Number of weighted sample cases	1,897,509	1,720,603	567,068	465,723

Table 8-18 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Level of distress in days following incident				
Not at all distressed	23.1 %	15.7 %	33.0 %	39.9 %
Mildly distressed	34.9	43.6	37.6	38.5
Moderately distressed	23.0	24.0	23.4 †	14.2
Severely distressed	19.0	16.6	6.0	7.4
Number of weighted sample cases	314,046	320,180	1,016,394	934,701

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

When asked about their level of distress at the time of the interview, the levels go down significantly. Whereas at the time of the incident, respondents expressed moderate or severe distress in approximately 4 in 10 incidents, these levels drop to approximately one in four incidents when looking back on the incident (table 8-19). Again, there was a strong relationship to the type of incident. For victims of rape, roughly one-third of respondents are still at least moderately distressed when thinking about the incident. This percentage goes down slightly for sexual assault and drops even further for the other incidents.

Table 8-19. Level of distress at the time of interview, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Level of distress at the time of interview				
Not at all distressed	30.8 %†	44.8 %	19.4 %	21.8 %
Mildly distressed	42.9 †	32.8	41.2	45.7
Moderately distressed	20.9	15.5	28.6	21.1
Severely distressed	5.4	6.9	10.8	11.5
Number of weighted sample cases	1,083,563	1,076,146	394,372	284,372

Table 8-19 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Level of distress at the time of interview				
Not at all distressed	24.0 %	41.2 %	41.4 %†	57.1 %
Mildly distressed	43.5	36.4	44 †	25.3
Moderately distressed	31.5 †	12.4	11.9	13.9
Severely distressed	1.0 †!	10.0	2.7	3.6
Number of weighted sample cases	161,152	201,681	528,039	590,092

Note: Item was only asked about the first incident in the detailed incident form. Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.12 Emotions Felt as Result of the Incident

Respondents who felt mildly, moderately, or severely distressed after the incident were asked whether they felt particular types of emotions after the incident for less than one month, for one month or longer, or not at all (table 8-20). Rape victims were more likely to have felt each of the emotions for one month or longer when compared to those experiencing sexual assault or the other incidents. Fully half or more of rape victims reported feeling worried or anxious, angry, sad or depressed, vulnerable, violated, and like they couldn't trust people for at least one month. This is higher for victims of sexual assault, which in turn is higher than for victims of other types of incidents. There were no differences by mode of interview among rape incidents.

Table 8-20. Emotions felt as result of the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Emotions as a result of the incident				
Worried or anxious				
Did not feel this way	30.9 %	33.9 %	13.8 %	18.2 %
Less than one month	28.8	31.6	28.1	32.9
One month or more	40.3	34.4	58.1	48.9
Angry				
Did not feel this way	20.0 %†	31.9 %	15.0 %	18.7 %
Less than one month	37.8	34.0	30.4	28.6
One month or more	42.2	34.1	54.5	52.7
Sad or depressed				
Did not feel this way	43.4 %†	55.5 %	28.5 %	27.1 %
Less than one month	23.5	17.9	18.1	26.3
One month or more	33.1	26.6	53.4	46.6
Vulnerable				
Did not feel this way	30.1 %	32.3 %	8.9 %	15.5 %
Less than one month	31.9	30.3	29.9	32.9
One month or more	38.1	37.4	61.3	51.6
Violated				
Did not feel this way	15.9 %	11.7 %	6.7 %	2.3 %!
Less than one month	38.0	41.2	36.0	29.8
One month or more	46.1	47.0	57.3	67.8
Like you couldn't trust people				
Did not feel this way	34.6 %	36.1 %	20.0 %	24.9 %
Less than one month	17.6	18.6	14.5	18.9
One month or more	47.9	45.3	65.5	56.2
Fearful				
Did not feel this way	47.5 %	53.9 %	24.8 %	40.6 %
Less than one month	24.9	20.5	29.4	23.4
One month or more	27.6	25.5	45.8	36.0
Number of weighted sample cases	1,275,732	1,280,324	479,319	411,832

Table 8-20. Emotions felt as result of the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015^v (continued)

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Emotions as a result of the incident				
Worried or anxious				
Did not feel this way	34.7 %	36.1 %	43.6 %	43.9 %
Less than one month	19.5	31.5	32.7	30.9
One month or more	45.8	32.4	23.7	25.3
Angry				
Did not feel this way	13.0 %†	41.3 %	26.5 %	36.7 %
Less than one month	41.5	28.1	42.6	40.4
One month or more	45.5	30.6	31.0	22.9
Sad or depressed				
Did not feel this way	39.5 %	59.0 %	57.1 %†	73.7 %
Less than one month	30.3	19.9	25.5 †	11.1
One month or more	30.2	21.0	17.4	15.3
Vulnerable				
Did not feel this way	39.4 %	30.6 %	44.0 %	44.2 %
Less than one month	22.7	27.3	36.8	29.9
One month or more	37.9	42.0	19.2	25.9
Violated				
Did not feel this way	16.8 %†	2.7 %!	23.1 %	22.3 %
Less than one month	35.2	48.8	40.6	45.7
One month or more	48.0	48.5	36.3	32.0
Like you couldn't trust people				
Did not feel this way	31.0 %	31.3 %	47.8 %	46.1 %
Less than one month	22.5	15.3	18.3	19.9
One month or more	46.5	53.5	33.9	34.0
Fearful				
Did not feel this way	65.7 %	49.0 %	59.5 %	65.4 %
Less than one month	13.9	19.4	25.1	19.0
One month or more	20.4	31.5	15.3	15.6
Number of weighted sample cases	209,967	272,168	586,447	596,325

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.13 Perceptions of the Incident as a Crime at the Time of the Incident

Respondents were asked whether they considered the incident to be a crime. As with the distress question, this was asked for at the time the incident occurred, as well as at the time of the interview. Less than half of the rape victims considered it a crime at the time of the incident (42%). This percentage drops for victims of sexual assault (14% for ACASI; 28% for CATI) and is about the same percentage for victims of unwanted contacts (table 8-21).

Table 8-21. Respondent considered incident to be a crime at the time, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent considered incident to be a crime at the time?				
Yes	19.7 %	25.9 %	42.0 %	42.6 %
Number of weighted sample cases	2,681,107	1,888,115	561,398	453,958

Table 8-21 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent considered incident to be a crime at the time?				
Yes	14.5 %	27.7 %	13.7 %	18.5 %
Number of weighted sample cases	362,323	331,900	1,757,386	1,102,257

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Respondents who said they did not consider the incident to be a crime were asked to provide reasons why they did not consider it to be a crime. Results differed significantly by mode, with CATI respondents generally endorsing more items than those on the ACASI. This is likely due to the way these items were presented for each mode. On the telephone, the interviewer read each

response and asked for a yes/no answer, while on the ACASI respondents were asked to “mark all that apply.”

For rape victims, some of the most common reasons were related to the victim’s perceptions related to the offender’s actions or state of mind, including the person stopped when the victim resisted (29.2% ACASI; 30.6% CATI) and the victim didn’t think the person knew the offender knew what she wanted to happen (20.4% ACASI; 39.0% CATI). The other more common reasons were she didn’t think it was against the law (32.5% ACASI; 57.3% CATI) and because the victim had been using alcohol/drugs at the time (23.6% ACASI; 89.0% CATI).

Among sexual assault incidents, the primary reasons for not considering the incident to be a crime were because the woman did not think it was against the law (66.9% CATI, 44.3% ACASI), and because the person stopped when she resisted (57.6% CATI, 37.0% ACASI) (table 8-22).

Table 8-22. Reasons why respondent did not consider incident to be a crime at the time, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Reasons why respondent did not consider incident to be a crime at the time				
Because the person stopped when I resisted	47.2 %†	61.3 %	29.2 %	30.6 %
Because I didn't think the person knew what I wanted to happen	15.3 %†	35.5 %	20.4 %	39.0 %
Because I didn't think it was against the law	35.1 %†	65.4 %	32.5 %†	57.3 %
Because I had been using alcohol/drugs at the time	10.9 %†	47.9 %	23.6 %†	89.0 %
Because offender was spouse/boyfriend/relative	3.5 %	5.0 %	6.5 %	16.2 %
Because offender was elderly/drunk/mental health	2.1 %	3.1 %	0.9 %!	1.4 %!
Some other reason	11.7 %	8.8 %	12.4 %	13.5 %
Number of weighted sample cases	2,152,685	1,398,624	325,804	260,513

Table 8-22 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Reasons why respondent did not consider incident to be a crime at the time				
Because the person stopped when I resisted	37.0 %†	57.6 %	53.3 %†	71.2 %
Because I didn't think the person knew what I wanted to happen	13.4 %†	31.0 %	14.6 %†	35.7 %
Because I didn't think it was against the law	44.3 %†	66.9 %	33.8 %†	67.3 %
Because I had been using alcohol/drugs at the time	17.0 %†	47.4 %	6.8 %†	32.8 %
Because offender was spouse/boyfriend/relative	2.1 %!	--	3.2 %	3.1 %
Because offender was elderly/drunk/mental health	3.0 %!	7.7 %	2.2 %	2.4 %
Some other reason	12.8 %	7.5 %	11.3 %	7.8 %
Number of weighted sample cases	309,899	240,065	1,516,982	898,046

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group. -- Less than 0.05%

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

When asked whether they now considered the incident to be a crime, the percentage went up. For example, in the general population, 67.2 percent and 77.2 percent of the rape victims asked through ACASI and CATI, respectively, thought it was a crime. This is an increase of 20 to 30 percentage points when comparing to what the victims thought at the time of the incident. There was a similar increase for sexual assaults and unwanted sexual contacts (table 8-23).

There were significant mode effects for each of the three types of incidents. ACASI respondents were less likely to say they thought it was a crime than did CATI respondents. For rape, there is a difference between modes of approximately 10 percentage points, for sexual assault of around 16 percentage points, and for unwanted contact of 6 percentage points.

Table 8-23. Respondent now considers incident to be a crime, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent now considers incident to be a crime?				
Yes	43.0 %†	53.1 %	67.2 %	77.2 %
Number of weighted sample cases	2,656,274	1,849,471	558,433	444,373

Table 8-23 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent now considers incident to be a crime?				
Yes	44.3 %	60.6 %	35.0 %	41.0 %
Number of weighted sample cases	362,323	325,177	1,735,518	1,079,921

Note: Estimates are based on weighted data. See Appendix A for standard errors.

† Significant difference from CATI at the 95% confidence level.

* Comparison group.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.14 Police Involvement

In less than 10 percent of incidents, respondents said that the police found out or were informed of the incident (table 8-24). Rape victims reported the highest numbers of the police being informed (roughly 15%) compared to other types of victimization (between 2.6% and 5.9%). There were no significant mode differences.

Table 8-24. Whether police were informed about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Were police informed about the incident?				
Yes	6.5 %	7.2 %	15.4 %	14.7 %
Number of weighted sample cases	2,674,737	1,904,290	567,468	455,813

Table 8-24 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Were police informed about the incident?				
Yes	2.6 %!	5.9 %	4.4 %	4.6 %
Number of weighted sample cases	362,323	336,736	1,744,947	1,111,741

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Respondents who considered the incident to be a crime at the time, but did not report it to the police, were asked why they did not report it to the police (table 8-25). The sample sizes for these tabulations are relatively small and have very high relative standard errors, especially for sexual assault and other incidents. In terms of patterns, among rape victims the primary reasons for not reporting were related to fear of anyone else finding out (“I did not want anyone else to know”) and fear of retaliation, as well as a feeling it would be hard to prove the case (“It was my word against the person’s”). For sexual assault and the other types of incidents, the most frequent answer was that the incident was not serious enough to report.

There were no significant differences by mode.

Table 8-25. Reasons why police were not informed about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Reasons why police were not informed about the incident				
I did not want anyone else to know	45.8 %	42.6 %	64.8 %	58.1 %
It was my word against the person's	30.3	43.7	48.8	46.9
The police wouldn't think it was a crime	16.1 †	52.5	21.6	47.5
I was afraid of being treated with hostility by police or lawyers	17.0	32.2	22.1 !	48.5
I was afraid of retaliation by the person or others	27.7 †	49.6	46.4	61.0
I did not think it was serious enough to report	40.8 †	64.6	28.0	52.7
Some other reason	18.7	16.5	15.6	26.1
Number of weighted sample cases	407,431	387,199	168,842	139,722

Table 8-25 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Reasons why police were not informed about the incident				
I did not want anyone else to know	29.8 %	38.1 %	33.0 %	31.8 %
It was my word against the person's	36.9	61.3	12.9 †	32.8
The police wouldn't think it was a crime	12.7 †!	57.4	12.1 †	54.5
I was afraid of being treated with hostility by police or lawyers	24.5 !	31.3	10.9 !	19.1
I was afraid of retaliation by the person or others	31.6	58.7	10.6 †	35.8
I did not think it was serious enough to report	63.0	56.6	46.9 †	79.0
Some other reason	17.4 !	29.2	21.6 †	2.4 !
Number of weighted sample cases	42,912	79,126	195,677	168,350

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group. Includes incidents respondents said was a crime.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.2.15 Involvement of Others

Respondents were asked about other individuals or agencies that might have been informed about the incident, including mental health professionals, crisis hotline operators, family and friends, and organizations providing assistance to women and victims.

Women were asked if they had spoken to a psychologist, psychiatrist, counselor, or other mental health professional about the incident. These items were asked only on the long version of the DIF, which was asked for the incident that had the highest priority in the DIF selection algorithm. Consequently, these items were administered only for a subset of the incidents reported, with rape being the most likely to be covered.³⁸ Roughly 10 to 12 percent of respondents did speak with a mental health professional after the incident occurred (table 8-26). In incidents involving rape, this rate was as high as 22.3 percent for ACASI respondents and 15.0 percent for CATI respondents, although differences between modes of interviewing were not significant. Very few respondents reached out to a crisis hotline operator (2.7% ACASI, 0.3% CATI), although this rate was higher for ACASI rape incidents (9.4%) (table 8-27).

³⁸This algorithm gave highest priority to rapes reported on the SV screener, then unwanted penetration or attempted penetration, and then unwanted sexual contact/non-contact.

Table 8-26. Whether respondent talked to a mental health professional, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent talked to mental health professional about incident?				
Yes	12.9 %	10.6 %	22.3 %	15.0 %
Number of weighted sample cases	1,453,247	1,074,662	394,372	281,072

Table 8-26 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent talked to mental health professional about incident?				
Yes	11.8 %	7.7 %!	9.0 %	9.4 %
Number of weighted sample cases	164,746	198,881	894,129	594,708

Note: Item was only asked about the first incident in the detailed incident form. Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-27. Whether respondent talked to a crisis hotline operator, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent talked to a crisis hotline operator about incident?				
Yes	2.7 %†	0.3 %!	9.4 %	--
Number of weighted sample cases	1,453,247	1,074,662	394,372	281,072

Table 8-27 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent talked to a crisis hotline operator about incident?				
Yes	1.3 %!	1.7 %!	-- %	--
Number of weighted sample cases	164,746	198,881	894,129	594,708

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The DIF included items asking if the victim had talked to someone else, besides the police. These items and all the items discussed below were asked for all DIFs filled out by the respondents, not just the first one. Aside from the police, more than half of the victims of unwanted contact told someone else about the incident, and two-thirds of rape victims told someone else (table 8-28). In most cases, this other person was a friend (77.7% ACASI, 88.3% CATI), but roughly one-third also told a spouse, boyfriend, or partner (31.4% ACASI, 39.3% CATI) (table 8-29). Among rape incidents, CATI respondents were slightly more likely to say yes to most of the categories, though this only reached levels of significance for telling parents about the incident (16.7% ACASI, 34.6% CATI). As discussed earlier, higher rates of endorsement on CATI are likely due to the way these items were presented for each mode. On the telephone, the interviewer read each response and asked for a yes/no answer, while on the ACASI respondents were asked to “mark all that apply.”

Table 8-28. Respondent told others about the incident, other than police, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent told others about the incident?				
Yes	56.0 %†	65.8 %	66.0 %	69.1 %
Number of weighted sample cases	2,675,645	1,904,340	567,468	462,422

Table 8-28 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent told others about the incident?				
Yes	54.9 %	62.3 %	53.0 %†	65.5 %
Number of weighted sample cases	362,323	336,736	1,745,854	1,105,182

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-29. Whom respondent told about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Whom respondent told about the incident				
Parents or a parent	18.7 %	23.9 %	16.7 %†	34.6 %
Husband, boyfriend, partner	31.4 %	39.3 %	29.1 %	44.2 %
Family member other than parents	18.9 %	27.5 %	17.9 %	33.6 %
A friend	77.7 %†	88.3 %	71.9 %	86.8 %
Co-worker/boss/manager/teacher/employer/HR	3.8 %	5.9 %	1.7 %!	1.6 %!
Someone else	5.3 %†	10.8 %	11.5 %	20.5 %
Number of weighted sample cases	1,497,934	1,252,675	374,303	319,305

Table 8-29 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Whom respondent told about the incident				
Parents or a parent	17.9 %	27.2 %	19.7 %	18.2 %
Husband, boyfriend, partner	31.6 %	47.6 %	32.2 %	34.7 %
Family member other than parents	20.0 %	30.8 %	19.0 %	23.9 %
A friend	88.3 %	77.6 %	77.7 %†	92.2 %
Co-worker/boss/manager/teacher/employer/HR	1.2 %!	2.4 %!	5.3 %	8.8 %
Someone else	4.7 %!	14.8 %	3.0 %	5.4 %
Number of weighted sample cases	198,925	209,719	924,706	723,651

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

In fewer than 5 percent of incidents, victims told an organization other than the police about the incident (table 8-30). This rate was slightly higher among rape incidents, with 14.2 percent of ACASI respondents and 7.3 percent of CATI respondents speaking with an organization about what happened. Among this small number of rape incidents, victims spoke primarily with a counselor or therapist who was unaffiliated with a rape crisis center or victim services hotline (79.3% ACASI, 85.6% CATI) (table 8-31).

Table 8-30. Whether respondent told any organization about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent told any organization about the incident, other than police?				
Yes	4.3 %	4.2 %	14.2 %	7.3 %
Number of weighted sample cases	2,682,903	1,908,908	567,468	462,422

Table 8-30 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent told any organization about the incident, other than police?				
Yes	2.4 %!	7.4 %!	1.5 %	1.9 %
Number of weighted sample cases	362,323	336,736	1,753,112	1,109,750

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-31. What organization respondent told about the incident, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	All sexual victimizations		Rape ^a	
	ACASI	CATI	ACASI	CATI
Type of organization respondent told about the incident				
Women's programs or service	19.4 %!	47.9 %	22.3 %!	s
Rape crisis center or victim services hotline	38.4 %	10.6 %!	41.9 %	s
Counselor or therapist not associated with a rape crisis center or victim services hotline	81.9 %	78.5 %	79.3 %	85.6 %
Other	18.0 %	37.7 %	16.7 %!	s
Number of weighted sample cases	115,536	79,299	80,452	33,680

Table 8-31 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI	CATI	ACASI	CATI
Type of organization respondent told about the incident				
Women's programs or service	s	s	s	s
Rape crisis center or victim services hotline	s	s	s	s
Counselor or therapist not associated with a rape crisis center or victim services hotline	s	s	s	s
Other	s	s	s	s
Number of weighted sample cases	8,819	24,752	26,265	20,867

Note: For women who contacted an organization besides the police. Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

s Data suppressed for disclosure reasons.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

When asked specifically whether they received any help from an office or agency (other than the police) that deals with victims of crime, only 2 to 3 percent said “yes” (table 8-32). Among rape victims, fewer than 10 percent sought out this type of help (8.4% ACASI, 6.2% CATI). Among the small number of incidents in which women sought out this type of help, respondents were primarily receiving legal advice, counseling, and help to remove themselves from danger. Few were seeking compensation for injury (table 8-33).

Table 8-32. Whether respondent received any help from an agency that deals with victims of crime, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent received help from any office or agency that deals with victims of crime, other than police?				
Yes	2.2 %	3.3 %	8.4 %	6.2 %
Number of weighted sample cases	2,674,742	1,908,908	564,033	462,422

Table 8-32 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Respondent received help from any office or agency that deals with victims of crime, other than police?				
Yes	--	3.0 %!	0.6 %!	2.2 %!
Number of weighted sample cases	362,323	336,736	1,748,386	1,109,750

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-33. Type of help received, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014–2015

	All sexual victimizations		Rape ^a	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Type of help respondent received from agency that deals with victims of crime				
Psychological or emotional counseling	56.8 %	88.3 %	52.2 %!	s
Compensation for your injury	3.4 %!	18.4 %!	4.2 %!	s
Help to remove you from danger	66.6 %	50.6 %	64.4 %	s
Legal advice or advocacy	66.7 %†	20.8 %	76.6 %	s
Number of weighted sample cases	58,301	63,549	47,360	28,873

Table 8-33 (continued)

	Sexual assault ^b		Other incidents ^c	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Type of help respondent received from agency that deals with victims of crime				
Psychological or emotional counseling	s	s	s	s
Compensation for your injury	s	s	s	s
Help to remove you from danger	s	s	s	s
Legal advice or advocacy	s	s	s	s
Number of weighted sample cases	.	10,099	10,942	24,577

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

s Data suppressed for disclosure reasons.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3 Comparing Circumstances for the RSA Pilot Test and NCVS

This section describes comparisons of the circumstances of the incidents measured on the NCVS and the RSA Pilot Test. For the NCVS, the results are based on NCVS files available through the Inter-University Consortium for Political and Social Research (ICPSR). The data presented below are for women ages 18 to 49 nationally who had reported a rape or sexual assault from 2005

to 2014. The national data were used because there were not enough incidents to analyze when concentrating on the five CBSAs in the study.

The comparisons discussed below are for all rapes and sexual assaults reported on each survey. For the RSA Pilot Test, data are presented for both modes and the total across modes. When there are not significant differences between the modes, the total is discussed.

In some instances, the wording of the questions on RSA Pilot Test and NCVS are not the same. If this is the case, the differences are described.

8.3.1 Time of Day of Incident

Table 8-34 shows a comparison of the time of day that the incident occurred. NCVS incidents were significantly more likely than RSA Pilot Test incidents to have taken place during the daytime (39.2% NCVS vs. 24.4% RSA Pilot Test). With respect to the specific time of day, the CATI results from the RSA Pilot Test tend to be more comparable to the NCVS.³⁹ The NCVS victims identified more incidents occurring in the afternoon and fewer in the middle of the night than on the RSA Pilot Test.

³⁹The ACASI offered an explicit “don’t know” category, which is why there are several of these “don’t know” responses for this mode.

Table 8-34. Time of day incident occurred, by study and mode of interview for females ages 18-49 in the general population

	NCVS ^{a*}	RSA Pilot Test ^b		
		TOTAL	ACASI ^c	CATI ^d
Time of day				
Daytime	39.2 %	24.4 %†	23.1 %†	25.9 %†
Between 6 a.m. and 12 noon	11.2 %	6.2 %	4.3 %†	8.4 %
Between 12 noon and 6 p.m.	22.1 %	12.8 %†	10.5 %†	15.4 %
Do not know time of day	5.9 %!	5.5 %	8.4 %	2.1 %
Nighttime	60.8 %	75.6 %†	76.9 %†	74.1 %†
Between 6 p.m. and 12 midnight	37.1 %	37.9 %	36.1 %	39.9 %
Between 12 midnight and 6 a.m.	19.7 %	29.0 %†	25.8 %	32.6 %†
Do not know time of night	4.0 %	8.8 %†	14.9 %†	1.6 %†
Number of weighted sample cases	1,793,324	1,729,308	930,799	798,508

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^b Based on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015. .

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.2 Location of Incident

Both NCVS and RSA Pilot Test asked about the location of the incident; however, the RSA Pilot Test question was not asked exactly as it was on the NCVS. The NCVS asked respondents if the incident took place in or near your home; in, at, or near a friend/relative/neighbor's home; at a commercial place; in a parking lot or garage, at school, in an open space, or somewhere else. The RSA Pilot Test had three categories, "at home," "at someone else's home," and "somewhere else." This was created from the question that asked what the respondent was doing at the time of the incident. The equivalent categories were created using the NCVS location code.

As shown in Table 8-35, more than twice as many of the NCVS incidents took place at the respondent's home than are reported on the RSA Pilot Test (54.0% NCVS, 24.7% RSA Pilot Test ACASI, 25.4% RSA Pilot Test CATI). In contrast, roughly three times as many RSA Pilot Test

incidents took place at someone else's home than are reported on NCVS (10.5% NCVS, 32.6% RSA Pilot Test ACASI, 28.9% RSA Pilot Test CATI).

Table 8-35. Location at time of incident, by study and mode of interview for females ages 18-49 in the general population

Location at time of incident	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
At home	54.0 %	25.0 %†	24.7 %†	25.4 %†
At someone else's home	10.5	30.9 †	32.6 †	28.9 †
Somewhere else ^e	35.5	44.1 †	42.7	45.7 †
Number of weighted sample cases	1,918,332	1,797,198	944,355	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015..

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eIncludes all locations other than at or near respondent's home, or at or near another home.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.3 Use of a Weapon

NCVS and RSA Pilot Test both asked about the presence of a weapon during the incident. The questions were slightly different. On NCVS, the question is asked, "Did the offender have a weapon such as a gun or knife, or something to use as a weapon, such as a bottle or wrench?" On RSA Pilot Test, the question asked whether the person "used a weapon or threatened to use a weapon."

The results are similar between the two surveys. On the NCVS, 7.2 percent reported a weapon was present, compared to 8.8 percent for the RSA Pilot Test (table 8-36).

Table 8-36. Use of a weapon, by study and mode of interview for females ages 18-49 in the general population

	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Use of a weapon during incident?				
Yes	7.2 %	8.8 %	9.8 %	7.8 %
Number of weighted sample cases	1,803,165	1,784,629	931,785	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015..

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.4 Injury from Incident

NCVS and RSA Pilot Test both ask about injuries sustained during the incident, but in different ways. On the NCVS, the interviewer asks an open-ended question “what were the injuries suffered, if any?” (Q31a NCVS-2). The interviewer codes the response into one of the following categories:

1. None
2. Raped
3. Attempted rape
4. Sexual assault other than rape or attempted rape
5. Knife or stab wounds
6. Gunshot, bullet wounds
7. Broken bones or teeth knocked out
8. Internal injuries
9. Knocked unconscious
10. Bruises, black eye, cuts, scratches, swelling, chipped teeth
11. Other.

The coding of this is complicated. The “raped” code is automatically assigned to anyone who had previously reported rape when they were asked about the method of attack in a prior question (Q29a on NCVS-2). It is also coded this way if a rape is reported as an injury, but rape had not been reported at any other point on the detailed incident form.⁴⁰ Attempted rape is coded only when this was previously reported as a method of attack and another physical injury (i.e., codes 5-11 above) is reported.

On the RSA Pilot Test, the respondent was first asked if any injuries occurred (yes/no) (table 8-37). If this was a “yes,” then then a list was presented that contained the same codes as 5-11 for the NCVS question with two exceptions—one, the RSA Pilot Test also included a category “injury from sexual intercourse, such as vaginal or anal tearing,” and two, the RSA Pilot Test combined gunshot and knife wounds into a single category.

Given the differences between the coding of these items, comparisons are restricted to the comparable physical injury items. Rates of injuries were higher on the NCVS than the RSA Pilot Test (32.4% for NCVS; 20.1% for RSA Pilot Test) (table 8-37). With respect to the injuries that are comparable across the surveys, the NCVS had higher rates of bruises, black eyes, cuts, scratches, swelling, and chipped teeth (29.8% for NCVS; 16.7% for RSA Pilot Test) (table 8-38). Very few of the victims reported other types of injuries, such as gunshot or stab wounds, internal injuries, or being knocked unconscious.

⁴⁰If the respondent reports rape for the first time at the injury question, the interviewer probes whether she means forced or coerced sexual intercourse. It is coded as an injury if the respondent confirms that it was forced or coerced sexual intercourse.

Table 8-37. Physical injuries, by mode of interview for females ages 18-49 in the general population

Any physical injuries as a result of the incident?	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Yes ^e	32.4 %	20.1 %†	22.0 %	17.9 %†
Number of weighted sample cases	1,918,332	1,802,954	950,110	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015. .

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eNCVS data reflect those who said yes to one or more type of physical injury.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-38. Type of physical injuries, by mode of interview for females ages 18-49 in the general population

Type of physical injuries	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Injury from sexual intercourse ^e	~	7.1 %†	9.8 %†	4.2 %†
Gunshot or stab wounds ^f	0.4	0.8 !	1.1 !	0.4 !
Broken bones or teeth knocked out	2.9 !	1.5	1.7 !	1.2 !
Bruises, black-eye, cuts, scratches, swelling, chipped teeth	29.8	16.7 †	17.7 †	15.6 †
Internal injuries	5.7	1.9	3.5	-- †
Knocked unconscious	2.3	2.0	3.2	0.6 †!
Other injuries	1.9	0.9	0.7 !	1.2
Number of weighted sample cases	1,918,332	1,809,153	956,310	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

~ Not offered on the NCVS.

† Significant difference from NCVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015. .

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eRSA Pilot Test included a category of injury from sexual intercourse, such as to your vagina or anus, which was not included in NCVS.

^fNCVS coded knife/stab wounds and gunshot/bullet wounds separately.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Slightly more victims on the NCVS reported receiving medical care than the victims on the RSA Pilot Test (table 8-39; 18.9% NCVS, 12.6% ACASI, 11.7% CATI). The NCVS respondents were more than twice as likely as RSA Pilot Test respondents to report medical care being received at a medical facility, though the categories on NCVS have a wider definition than on the RSA Pilot Test⁴¹ (table 8-40; 15.2% NCVS, 7.4% RSA Pilot Test ACASI, 5.6% RSA Pilot Test CATI). Overnight hospital stays were comparable between the NCVS and RSA Pilot Test incidents (table 8-41; 1.8% NCVS, 3.6% RSA Pilot Test ACASI, 1.9% RSA Pilot Test CATI).

Table 8-39. Medical care, by mode of interview for females ages 18-49 in the general population

Any medical care received as a result of the incident?	NCVS ^{**}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Yes	18.9 %	12.2 %	12.6 %†	11.7 %
Number of weighted sample cases	1,918,332	1,809,153	956,310	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015. .

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

⁴¹The NCVS also included care received at a health unit at work or school, or health care received at a first aid station. The RSA Pilot Test only included medical settings such as a doctor's office, emergency room, or hospital

Table 8-40. Location of medical care, by mode of interview for females ages 18-49 in the general population

Location where medical care was received	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
At the scene	3.2 %!	2.1 %	1.3 %!	2.9 %!
At home, at a neighbors or a friends	3.2	7.5 †	6.6	8.5 †
At a medical setting such as an emergency room, hospital, clinic, or doctor's office ^e	15.2	6.5 †	7.4 †	5.6 †
Some other place	0.3 !	0.5 !	0.4 !	0.7 !
Number of weighted sample cases	1,354,183	1,809,153	956,310	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eNCVS includes care received at a health unit at work/school, doctor's office, health clinic, emergency room, or hospital other than emergency room. RSA includes doctor's office, emergency room, or hospital.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 8-41. Whether respondent stayed overnight in the hospital, by mode of interview for females ages 18-49 in the general population

Overnight stay in hospital?	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Yes	1.4 %	2.8 %	3.6 %!	1.9 %!
Number of weighted sample cases	1,918,332	1,809,153	956,310	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.5 Number of Offenders

Both NCVS and RSA Pilot Test ask respondents to report the number of people involved in the incident.

There were no significant differences between NCVS and RSA Pilot Test results, with roughly 9 out of 10 incidents being committed by a single offender (table 8-42).

Table 8-42. Number of offenders, by study and mode of interview for females ages 18-49 in the general population

	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Number of offenders				
One	94.8 %	91.7 %	88.3 %†	95.6 %
More than one	5.2 %	8.3 %	11.7 %†	4.4 %
Number of weighted sample cases	1,858,987	1,799,055	954,818	844,237

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^b Based on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.6 Sex of Single Offender

Among incidents in which there was a single offender, respondents were asked whether the offender was male or female.

Data are comparable between the two surveys. On the NCVS, 97.2 percent of incidents involve a male offender compared to 97.1 percent for the RSA Pilot Test (table 8-43).

Table 8-43. Sex of single offender, by study and mode of interview for females ages 18-49 in the general population

	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Sex of offender				
Male	97.2 %	97.1 %	97.3 %	96.9 %
Female	2.8 %	2.9 %	2.7 %!	3.1 %!
Number of weighted sample cases	1,746,936	1,644,802	837,991	806,811

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.7 Relationship to Offender

Respondents to both NCVS and RSA Pilot Test were asked if they knew or had seen the offender before, or if the offender was a stranger. Respondents in both surveys who knew the offender were asked how they knew the person, using the categories by sight only, casual acquaintance, or well known. Both surveys asked for the relationship of the offender if the person was a casual acquaintance or well known. The method of collecting this information differed somewhat by the two studies. NCVS asks an open-ended question: "How did you know the offender? For example, was the offender a friend, cousin, etc.?" In contrast, the RSA Pilot Test has follow-up questions to determine if the person was a relative or not, and then asks for more details about the type of relationship.

To compare the surveys to each other, we applied the seven-category typology referenced earlier in this chapter to both studies (table 8-44). There were approximately the same percentage of incidents occurring by a casual acquaintance (9.0% NCVS; 11.8% RSA Pilot Test) or by someone known by sight only (5.5% NCVS; 6.2% RSA Pilot Test). However, the two surveys differed in the percentage of incidents perpetrated by strangers (14.6% NCVS; 24.3% RSA Pilot Test). Conversely, the NCVS had higher rates than the RSA Pilot Test of incidents being committed by a spouse or ex-spouse (19.3% NCVS; 7.8% RSA Pilot Test) or a well-known person other than a friend or intimate

partner (13.9% NCVS; 5.0% RSA Pilot Test). In contrast, NCVS had lower rates of incidents being committed by a friend or ex-friend (13.3% NCVS; 28.7% RSA Pilot Test).

The percentage of intimate partners (spouse, ex-spouse, boyfriend/girlfriend/ex-boyfriend/ex-girlfriend) is lower for the RSA Pilot Test (21.3%) compared to the NCVS (39.2%), which is statistically different. This is due to the higher proportion of incidents involving spouses/ex-spouses on the NCVS relative to the RSA Pilot Test.

Table 8-44. The offenders relationship to the victim by study and mode of interview for females ages 18-49 in the general population

	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Offender relationship				
Spouse, ex-spouse ^e	19.3 %	7.8 %†	5.6 %†	10.1 %
Boy/girlfriend or ex-boy/girlfriend ^f	19.9 %	13.5 %	12.9 %	14.2 %
Friend or ex-friend ^g	13.3 %	28.7 %†	29.5 %†	28.0 %†
Other well-known person ^h	13.9 %	5.0 %†	6.5 %†	3.5 %†
Other relative ⁱ	4.5 %	2.6 %	3.5 %	1.8 %!
Casual acquaintance ^j	9.0 %	11.8 %	14.4 %	9.1 %
By sight only ^k	5.5 %	6.2 %	5.5 %	6.9 %
Stranger ^l	14.6 %	24.3 %†	22.1 %	26.5 %†
Number of weighted sample cases	1,708,891	1,625,872	822,647	803,226

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

^eIncludes those who indicated in NCVS or RSA Pilot Test that the person was "a spouse or ex-spouse." In RSA Pilot Test, also includes those who were father of the respondent's child.

^fIncludes those who indicated in NCVS or RSA Pilot Test that the person was a boyfriend, girlfriend, ex-boyfriend, or ex-girlfriend.

^gIncludes those who indicated in NCVS or RSA Pilot Test that the person was a friend or ex-friend.

^hIncludes those who indicated in NCVS or RSA Pilot Test that the person was well-known, but not a relative of the respondent.

ⁱIncludes those who indicated in NCVS or RSA Pilot Test that the person was a relative other than a spouse.

^jIncludes those who indicated in NCVS or RSA Pilot Test that the person was "a casual acquaintance."

^kIncludes those who indicated in NCVS or RSA Pilot Test that the person was not a stranger but was known "by sight only."

^lIncludes those who indicated in NCVS or RSA Pilot Test that the person was "a stranger you had never seen before."

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.8 Offender Use of Alcohol or Drugs

Respondents were asked whether they thought the offender had been using alcohol or drugs in the hours leading up to the incident. The NCVS question “Were any of the offenders drinking or on drugs, or don’t you know?” and responses were coded into yes/no/don’t know. Approximately 27.6 percent of the NCVS respondents said they did not know. If the NCVS respondent said “yes,” she was asked a follow-up question to determine “Which was it? (Drinking or on drugs),” to which the response options were drinking, on drugs, both, or could not tell which.

On the RSA Pilot Test, a single item was used, asking, “Had the person who did this to you been using alcohol or drugs in the hours leading up to the incident?” with response options of alcohol, drugs, both alcohol and drugs, and neither. No explicit option was given for “don’t know.” Consequently, the percentage that said “don’t know” on RSA Pilot Test was significantly lower than on the NCVS (3.6% ACASI and 10.0% CATI). Because of the large discrepancies in the percentages saying “don’t know,” this category is treated as missing data for the comparison below.

Once taking out the don’t know responses, results were somewhat different between the two studies. NCVS had higher rates of the offender not using any alcohol or drugs during the incident (45.7% NCVS; 32.1% RSA Pilot Test). RSA Pilot Test had higher rates of the offender using alcohol (30.7% NCVS; 42.0% RSA Pilot Test) or both alcohol and drugs (12.6% NCVS; 20.2% RSA Pilot Test), whereas NCVS had higher rates of the offender using drugs (11.0% NCVS; 5.8% RSA Pilot Test).

Table 8-45. Offender's use of alcohol or drugs in the hours leading up to incident, by study and mode of interview for females ages 18-49 in the general population

	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Offender's use of alcohol or drugs in the hours leading up to incident				
Alcohol	30.7 %	42.0 %†	42.2 %†	41.7 %
Drugs	11.0	5.8 †	5.5 †	6.1
Alcohol and drugs	12.6	20.2 †	21.1 †	19.0
Neither alcohol nor drugs	45.7	32.1 †	31.2 †	33.2 †
Number of weighted sample cases	1,239,669	1,665,444	917,523	747,921

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.9 Respondent Emotions After Incident

This series of questions collected information on the respondent's emotions following the incident. While the same emotions were asked about in both studies, there is an important difference in the way the questions were worded. NCVS asks respondents to report whether they felt a specific type of emotion for a month or longer. A list of six emotions is presented to the respondent. The RSA Pilot Test respondents were asked about the same list of emotions but whether they felt this way for one month or less, one month or longer, or not at all.

The data presented show the percentage incidents in which respondents felt any of these ways for one month or longer (table 8-46). In all cases, NCVS respondents were significantly higher for all emotions. While this may be true, it is possible that NCVS respondents may have been reporting on emotions they felt at any point after the incident, not just those they continued to feel for a month or more. If the responses on the RSA Pilot Test for less than one month and one month or longer are combined (data not shown), the RSA Pilot Test respondents are uniformly higher than the NCVS except on vulnerability.

Table 8-46. Emotions felt as result of the incident, by mode of interview for females ages 18-49 in the general population

Emotions felt for one month or longer as a result of the incident	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Worried or anxious	87.8 %	63.2 %†	65.7 %†	60.2 %†
Angry	73.0	60.2 †	61.7	58.5
Sad or depressed	73.5	54.9 †	57.9 †	51.3 †
Vulnerable	80.0	61.7 †	65.4 †	57.5 †
Violated	86.2	72.0 †	67.3 †	77.7
Like you couldn't trust people	84.1	69.3 †	69.4 †	69.1 †
Number of weighted sample cases	996,732	917,084	494,960	422,124

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^bBased on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.10 Police Informed About Incident

Respondents in both surveys were asked if the police were informed about the incident.

Question wording was identical in both instruments.

NCVS respondents were more than three times as likely to have informed the police about the incident (table 8-47).

Table 8-47. Police notification, by study and mode of interview for females ages 18-49 in the general population

	NCVS ^{a*}	RSA Pilot Test ^b		
		TOTAL	ACASI ^c	CATI ^d
Were police informed about the incident?				
Yes	34.4 %	10.7 %†	10.4 %†	11.0 %†
No	65.6	89.3 †	89.6 †	89.0 †
Number of weighted sample cases	1,910,519	1,722,340	929,791	792,549

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^aBased on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^b Based on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015..

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.3.11 Involvement of Others

When asked specifically whether they received any help from an office or agency (other than the police) that deals with victims of crime, the NCVS respondents were more than five times as likely to report receiving this type of help after the incident (26.0% NCVS, 5.0% RSA Pilot Test ACASI, 4.6% RSA Pilot Test CATI) (table 8-48).

Table 8-48. Whether respondent received any help from agency that deals with victims of crime by mode of interview for females ages 18-49 in the general population

Respondent received help from any office or agency that deals with victims of crime, other than police?	NCVS ^{a*}	RSA Pilot Test ^b		
		Total	ACASI ^c	CATI ^d
Yes	26.0 %	4.8 %†	5.0 %†	4.6 %†
Number of weighted sample cases	1,876,860	1,805,718	952,875	852,843

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from NCVS at the 95% confidence level.

^a Based on a concatenated incident file reflecting the years 2005-2014 among women ages 18-49 nationally who reported completed or attempted rape or sexual assault.

^b Based on incident-level data for females ages 18-49 in the general population sample who reported rape or sexual assault on the RSA Pilot Test, 2014-2015.

^cAudio computer-assisted self-interview.

^dComputer-assisted telephone interview.

Sources: Bureau of Justice Statistics, National Crime Victimization Survey, 2005-2014 and Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

8.4 Summary and Discussion

This chapter described the victims who reported a rape or sexual assault on the RSA Pilot Test, what types of incidents they reported, and how the incidents differ from those reported on the NCVS.

8.4.1 Who are the Victims?

Many of the personal characteristics examined were correlated with rape and sexual assault at a bivariate level. The strongest correlates associated with both types of victimization were age, marital status, and race/Hispanic origin. These were significant in both the bivariate and multivariate logistic regression models. The very youngest among the respondents, those ages 18-20, had a victimization rate for rape that was about three times as high as those just a few years older (25-29). Similarly, marital status was strongly related to both types of victimizations. Again for rape, those who are married had significantly lower rates than those who were not married. Those who seemed to have the highest risk were those who had been in a romantic relationship in the past year, but were not cohabitating with anyone. With respect to race/Hispanic origin, non-Hispanic whites have the highest rates for both rape and sexual assault. For rape, Hispanics had the lowest rates. The other variable associated with rape victimization in the summary model was income. Those in the lowest income group had the highest victimization rates.

These results are very similar to what has been found for the NCVS. In particular, analysis of the NCVS, both in bivariate (Planty et al., 2013) and multivariate analyses (Lauritsen, 2012), have found the same effects of age, marital status, race/Hispanic origin, and income. Similarly, the null effect associated with college enrollment for rape is consistent with a similar analysis using NCVS data (Sinzit & Langton, 2014).

In several instances, the bivariate results did differ by mode of interview. Several factors related to wealth, including income and home ownership, were related to rape victimization for ACASI but not for CATI. For income, the rates for the ACASI found those with the lowest incomes had the highest risk. Similarly, for ACASI, those who were renting their home had the highest risk. This was not the case for CATI. In both of these instances the ACASI is more consistent with what one would expect from extant research, as well as what has been found on the NCVS. One possible explanation for this difference may be nonresponse bias. It may be that those

in the lowest income groups who could not be reached by phone, but could be reached in person, differed with respect to their risk of victimization. Chapter 10 did not find strong evidence of nonresponse bias, but these analyses used a different set of outcomes (benchmarks to demographics, level of effort).

A second possible explanation is that the mode of communication (written vs. verbal) resulted in different measurements. For example, lower income individuals are likely to be less literate and less adept at working with written, rather than spoken, communication. This may have led to interpreting and answering the victimization questions on the ACASI somewhat differently than on the telephone. If this were true, however, one would also have expected a difference in the relationship of education with victimization between modes. This was not the case. Education was not related to rape at all for both modes. It was related to sexual assault, but this was the case for both modes.

8.4.2 What are the Circumstances of Victimization for the RSA Pilot Test?

The RSA Pilot Test collected a wide variety of information about the circumstances and nature of the rape and sexual assault incidents. The primary tactic used in rape incidents was being held or pinned down by the offender. This was also the most common type of physical force used for sexual assault, but to a lesser degree. A second tactic used for assault was grabbing, pushing, or pulling the victim. Injuries were incurred by about one-third of the rape victims, with more than half of those injured receiving some type of medical care.

The RSA Pilot Test collected information on the circumstances surrounding incidents related to alcohol and drug use. Determining whether someone is unable to consent is difficult to measure on a survey. The RSA Pilot Test collected several measures that provide a few more details on the respondent's condition at the time of the incident. According to these measures, a significant percentage of these incidents involve some type of alcohol use by the victim. Approximately 40 percent of rapes and sexual assaults involved the victim using alcohol or drugs in the hours leading up to the incident. Victims were classified as unable to consent if they either said they were unable to consent because of the substance or they indicated they were passed out for at least parts of the incidents. Approximately 23 percent of the incidents of rape and sexual assault involved a victim who met these criteria. In addition, around 10 percent of the victims of rape and sexual assault reported indications of being intoxicated but still able to consent (e.g., unable to walk or communicate, unable to remember the incident, made decisions they would not normally make).

Among those who were considered unable to consent, about half said they were unable to consent because of the alcohol/drugs. The rest said they were unconscious for at least part of the incident. Among those considered unable to consent, almost 70 percent showed signs they were intoxicated (unable to walk on their own, unable to communicate clearly), 80 percent said they were less able to physically resist, and 80 percent said they made decisions they otherwise would not make. About 40 percent said they thought the perpetrator was giving them drinks after they were clearly drunk.

The emotional consequences of the incidents related to the victimization were significant, especially for rape. At least half of the rape victims reported feeling angry, sad or depressed, vulnerable, and unable to trust other people for one month or more. Almost half reported being fearful for one month or more.

About 40 percent of the victims of rape considered the incident to be a crime at the time the incident occurred. The reasons why they did not consider it a crime included the person stopped when they resisted, they did not think the offender knew what they wanted to happen, or they just did not think it was against the law. Upon thinking of the totality of the circumstances, and perhaps after going through the RSA Pilot Test interview or thinking about it more over time, a significant number of victims change their view of the incident. At the time of the interview, about 70 percent said it was a crime.

About 15 percent of the rape victims said the police found out about the incident. This relatively low percentage reflects the large percentage of individuals who did not consider the incident to be a crime at the time it occurred. Other reasons for not reporting it to the police were fear of retaliation, fear of someone else finding out, or deciding that it was just their word against the offender's.

8.4.3 How do the Incidents on RSA Pilot Test Compare to Those on the NCVS?

The large difference in victimization rates raises the question of how the incidents collected on the RSA Pilot Test and NCVS differ. The estimates for rape, for example, differ by a factor of 40 to 50. Is the RSA Pilot Test measuring completely different types of incidents? Or are the incidents just more of the same type? Along some key dimensions, the incidents seem to be very similar. They

infrequently involve weapons (around 8% of the time). However, they involve one male offender, have similar rates of requiring medical care and involve the same percentage of offenders who are boyfriends/girlfriends, casual acquaintances or known by sight only.

The correlates of victimization risk were also very similar for the two studies. For the RSA Pilot Test, the strongest correlates associated with both rape and sexual assault were age, marital status, and race/Hispanic origin. Females ages 18 to 24 had victimization rates for rape that were significantly higher than those just a few years older (25-29); this trend continued into the older age groups. Those who are married had significantly lower rates than those not married. With respect to race/Hispanic origin, non-Hispanic white women have the highest rates. Household income was also significantly related to rates of rape. Those in the lowest income group had the highest victimization rates. Analysis of the NCVS, both in bivariate (Planty et al., 2013) and multivariate analyses (Lauritsen, 2012), have found the same effects of age, marital status, race/Hispanic origin, and income.

A second similar finding is that women enrolled in college did not exhibit higher rates of victimization than non-college students (Chapter 8). Nationally, there is considerable concern over the high rates of sexual violence among college students, as revealed by recent campus climate surveys (White House Task Force Report, 2014). In analysis of the NCVS, Sinzit and Langton (2014) found that, after controlling for age, women age 18 to 24 who are not in college have slightly higher rates than those who are in college. The RSA Pilot Test found that college enrollment did not increase risk. Those currently in college had a similar victimization rate as those who were not in college.

Despite these similarities, there were several areas where the types of incidents reported on the two surveys diverged. The largest difference between the two surveys was for the percentage of crimes reported to the police. The victims identified on the NCVS were more than three times as likely to report it to the police than those on the RSA Pilot Test, and were five times as likely to have received help from victim services after the incident. Almost two-thirds of the RSA Pilot Test victims said that at the time of the incident they did not consider the incident to be a crime. Many of the victims said they did not think it was against the law, perhaps because of their own perceptions of the dynamics of what happened (e.g., the perceived intentions of the offender; use of alcohol). While there is no comparable measure on the NCVS, it is probably safe to assume that most respondents who report an incident on the survey believe it was a crime of some type.

Related to the above difference, the NCVS incidents involved more serious physical and emotional consequences than the RSA Pilot Test. These types of incidents are more likely to be reported to the police. For example, although the questions between the RSA Pilot Test and NCVS are not entirely comparable, rates of injury on the NCVS were higher than those on the RSA Pilot Test. As such, the studies also differed on the type of medical attention received after the incident. The victims on the NCVS reported a greater proportion got medical attention in a hospital setting. NCVS respondents also seemed to report more emotional consequences, although the two measures were difficult to compare.

A third difference was how well the victim knew the offender. The RSA Pilot Test incidents were less likely to involve spouse or ex-spouse, and were more likely to involve a friend or ex-friend or a stranger. As a result, fewer of the offenders on the RSA Pilot Test are considered intimate partners by BJS's definition (e.g., 21.3% vs. 39.2% are intimate partners on RSA Pilot Test and NCVS, respectively). Related to this, a higher percentage of incidents on the NCVS were reported to occur at the victim's home, while a larger percentage on the RSA Pilot Test occurred in someone else's home.

There were two other characteristics that the two surveys differed. One was the role of alcohol and drugs by the offender, although the questions on the two surveys were not entirely comparable. Offenders in the NCVS were more likely to have not used any alcohol or drugs during the incident, whereas offenders in the RSA Pilot Test were more likely to have used alcohol or a combination of alcohol and drugs. The second area was the time of day the incident occurred. The NCVS incidents were more likely to occur during the day, while the RSA Pilot Test incidents were more likely to occur at night.

Overall, the above comparisons indicate there are some similarities between what is reported on RSA Pilot Test and the NCVS. However, there are a number of differences. The more intensive focus on sexual violence on the RSA Pilot Test screener seems to produce more incidents that may not be captured on the NCVS. The higher rates of reporting to the police on the NCVS is indicative of the incidents being more salient and framed in a context that respondents are most likely to report when asked the two or three questions about rape and sexual assault on the NCVS screening instrument. The larger percentage of incidents occurring among friends on the RSA Pilot Test, as well as the incidents that do not involve as many injuries or medical attention, may be indicative of events that are not as readily recalled on the current NCVS screening instrument, which has a focus on crime (e.g., using terms such as 'rape' and 'assault').

9. Assessment of Two-Stage Design for Measuring RSA

The RSA Pilot Test combined behaviorally specific questions (BSQs) with a detailed incident form (DIF). This design has several advantages, among which is that it collects several different measures of sexual victimization. One is collected on the SV screener and the second on the DIF. This chapter discusses features of each stage of the RSA Pilot Test design and then uses this design to assess the quality of the reports at each stage.

9.1 Sexual Victimization Screener

The sexual victimization screener (SV screener) asked respondents about 14 different types of unwanted sexual experiences, and whether they have experienced any of them within the past 12 months. The SV screener asked questions about rape first, followed by other types of unwanted penetration, unwanted sexual contact and sexual non-contact (e.g., flashing). Each time that a respondent said “yes” to an item in the SV screener, they were immediately asked a short series of follow-up questions to determine eligibility. Follow-up questions were asked for up to four incidents per screener item.

This section describes the three operational features of the screener: (1) the extent to which incidents were found to be ineligible, (2) the dating of the incidents and (3) the amount of missing data, including that due to capping the number of incidents for which dates were collected.

9.1.1 Incident Eligibility and Reports of Duplicates

The follow-up questions in the SV screener were used to determine whether an incident was eligible to receive a detailed incident form (DIF). Incidents were considered eligible if (1) the respondent confirmed that it happened within the past 12 months, and (2) it was a unique incident, and not part of a previously reported incident.

The analyses below tabulate the data by grouping the SV items into three incident types: (1) rape items (SV1-SV5), (2) other unwanted penetration (SV6-SV8) and (3) unwanted sexual acts (SV9-SV14).

General Population

Overall, approximately 70 percent (69.8% for ACASI, 72.8% for CATI) of incidents that were reported in the SV screener were determined to be eligible (table 9-1). This means that after asking follow-up questions, about 30 percent of all the incidents initially reported in the SV screener were not unique incidents.

The most common reason that an incident was not counted as eligible was when the respondent indicated that the incident actually occurred as part of another incident previously reported on the screener. For example, a respondent who experienced a single incident involving both forced rape and unwanted kissing might report this same incident in both items SV1 and SV9. When answering SV9, if the respondent indicated that the unwanted kissing occurred in the same month as the forced rape, she was asked whether the unwanted kissing was part of the same incident as the previously reported forced rape. Overall, just over 20 percent (20.7% for ACASI, 22.2% for CATI) of incidents reported in the SV screener turned out to be a duplicate of a previously reported incident. There does not seem to be a difference in the extent to which incidents were found to be duplicates across the two modes.

Table 9-1. Proportion of unique in scope incidents reported in victimization screener based on screener items, by mode of interview for females ages 18-49 in the general population, 2014-2015

	Number of weighted incidents		Eligible incidents ^a		Part of previous incident ^b	
	ACASI ^f	CATI ^g	ACASI ^f	CATI ^{g*}	ACASI ^f	CATI ^{g*}
Rate across all screening items ^h	5,501,872	4,642,144	69.8 %	72.8 %	20.7 %	22.2 %
Rape ⁱ	874,577	550,295	62.1 %†	77.6 %	25.0 %	15.9 %
SV 1: Forced vaginal sex	204,836	145,924	69.7	78.1	7.6 !	12.0 !
SV 2: Forced oral sex	136,666	107,201	58.1	83.1	25.2	16.9 !
SV 3: Forced anal sex	68,490	55,620	48.3	58.6	46.2	12.9 !
SV 4: Forced digital penetration	193,257	50,674	37.8	66.0	56.1	34.0 !
SV 5: Penetration while unable to consent	271,327	190,876	79.2	82.7	10.5	14.5
Other unwanted sex ^j	1,052,410	673,887	73.0 %	80.3 %	14.4 %	13.6 %
SV 6: Sex by coercion	94,299	83,047	59.5	79.6	34.1	20.4 !
SV 7: Other unwanted sex	134,231	54,600	79.7	66.8	8.4 !	29.5 !
SV 8: Attempted unwanted sex	823,880	536,241	73.4	81.7	13.1	10.9
Sexual contact ^k	3,574,885	3,417,962	70.7 %	70.5 %	21.5 %	24.9 %
SV 9: Unwanted kissing	529,112	433,157	72.8	73.6	19.4	19.5
SV10: Unwanted groping	1,525,227	1,315,351	78.3	74.0	14.2 †	22.0
SV11: Attempted unwanted kissing or groping	1,027,105	1,239,118	59.1	69.4	33.3	26.5
SV12: Flashing/unwanted exposure	289,895	252,600	78.1	68.5	12.1	25.2
SV13: Made you show your body parts	78,139	97,691	37.2	25.1	46.8	67.2
SV14: Unwanted sexual pictures or movies	125,408	80,046	69.5	74.6	28.7	25.4

Table 9-1. Proportion of unique in scope incidents reported in victimization screener based on screener items, by mode of interview for females ages 18–49 in the general population, 2014–2015 (continued)

	Out of scope ^c		Missing date information ^d	
	ACASI ^f	CATI ^{g*}	ACASI ^f	CATI ^{g*}
Rate across all screening items ^h	7.6 %†	3.6 %	1.9 %	1.5 %
Rape ⁱ	10.9 %	4.2 %	2.1 %!	2.3 %!
SV 1: Forced vaginal sex	17.8	6.5 !	4.9 !	3.4 !
SV 2: Forced oral sex	10.9 !	--	5.8 !	--
SV 3: Forced anal sex	5.6 !	21.2 !	--	7.4 !
SV 4: Forced digital penetration	6.1 !	--	--	--
SV 5: Penetration while unable to consent	10.4	1.1 !	--	1.7 !
Other unwanted sex ^j	9.9 %†	4.3 %	2.7 %!	1.8 %!
SV 6: Sex by coercion	6.4 !	--	--	--
SV 7: Other unwanted sex	9.6 !	3.7 !	2.2 !	--
SV 8: Attempted unwanted sex	10.3	5.1	3.1 !	2.3 !
Sexual contact ^k	6.2 %†	3.3 %	1.6 %	1.3 %
SV 9: Unwanted kissing	4.6	5.5	3.2 !	1.4 !
SV10: Unwanted groping	6.9 †	2.8	0.7 !	1.3
SV11: Attempted unwanted kissing or groping	5.7	2.8	1.9 !	1.3
SV12: Flashing/unwanted exposure	6.1	4.5 !	3.8 !	1.8 !
SV13: Made you show your body parts	16.0 !	7.7 !	--	--
SV14: Unwanted sexual pictures or movies	1.8 !	--	--	--

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

^aIncludes incidents that happened within the 12 months prior to the interview and were eligible for a detailed incident form.

^bIncludes incidents that the respondent indicated happened as part of another incident already mentioned.

^cIncludes incidents that happened outside of the 12-month reference period.

^dIncludes incidents for which respondents could not recall the date and could not confirm whether it happened during the 12-month reference period.

^eRefers to the screening item number and the content of the item for each sexual victimization screener item. See Appendix B for full question wording.

^fAudio computer-assisted self-interview.

^gComputer-assisted telephone interview.

^hCombined total across all screening items.

ⁱIncludes screening items measuring penetrative behaviors using force.

^jIncludes screening items measuring penetration using coercion, attempted penetration, and other unwanted penetration.

^kIncludes non-penetrative behaviors using any tactic.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

The second reason that an incident was determined to not be eligible was if the respondent indicated that the incident occurred prior to the beginning of the reference period. As noted above, all respondents were asked to date the incident with respect to the month and year of occurrence. If the respondent indicated that the incident happened more than 12 months ago, it was considered ineligible. If the respondent could not provide a month, the respondent was asked to confirm whether or not the incident occurred within the past 12 months. If they said that it had not, then the incident was counted as out of scope. In the general population, the overall proportion of SV screener incidents that were determined to have occurred more than 12 months ago was significantly higher for ACASI respondents (7.6%) than for CATI respondents (3.6%). For most of the SV screener items, there was not a statistically significant difference between modes. The pattern is generally for the ACASI to have a higher percentage of incidents dated as out of the reference period. For example, for penetration while unable to consent, 10.4 percent of the incidents were dated this way compared to 1.1 percent for the CATI. A significant difference in this same direction was also found for unwanted groping (6.9% vs. 2.8%).

Finally, an incident was not counted as eligible if a respondent said that they did not know or declined to answer the follow-up question asking whether or not the incident happened within the past 12 months. These incidents were counted as “missing date information” and account for fewer than 2 percent of all SV screener incidents (1.9% for ACASI, 1.5% for CATI).

Volunteer Sample

Among the volunteer sample (table 9-2), just over 73 percent of incidents reported in the SV screener were determined to be eligible (74.2% for ACASI, 73.3% for CATI). Generally, the proportion of incidents that were eligible tended to be higher for incidents involving penetration (approximately 80 percent eligible) than for incidents involving sexual touching (approximately 70 percent eligible).

Table 9-2. Proportion of unique in scope incidents reported in victimization screener based on screener items, by mode of interview for females ages 18-49 in the volunteer sample, 2014-2015

Sexual victimization screener: Item number and description ^e	Number of unweighted incidents		Eligible incidents ^a		Part of previous incident ^b	
	ACASI ^f	CATI ^g	ACASI ^f	CATI ^{g*}	ACASI ^f	CATI ^{g*}
Rate across all screening items ^h	2,446	1,487	74.2 %	73.3 %	19.5 %†	23.3 %
Rape ⁱ	386	177	77.7 %	82.5 %	17.6 %	14.1 %
SV 1: Forced vaginal sex	96	49	92.7	93.9	2.1 !	2.0 !
SV 2: Forced oral sex	75	24	60.0	66.7	38.7	29.2
SV 3: Forced anal sex	23	14	82.6	92.9	17.4	7.1 !
SV 4: Forced digital penetration	49	29	57.1	69.0	36.7	24.1
SV 5: Penetration while unable to consent	143	61	83.2	83.6	10.5	14.8
Other unwanted sex ^j	516	224	80.6 %	76.3 %	9.1 %†	21.0 %
SV 6: Sex by coercion	61	46	73.8	60.9	13.1 †	34.8
SV 7: Other unwanted sex	23	26	78.3	76.9	8.7 !	19.2
SV 8: Attempted unwanted sex	432	152	81.7	80.9	8.6 †	17.1
Sexual contact ^k	1,544	1,086	71.1 %	71.2 %	23.5 %	25.2 %
SV 9: Unwanted kissing	240	186	66.7	65.1	27.5	31.2
SV10: Unwanted groping	644	422	77.2	76.5	18.5	20.1
SV11: Attempted unwanted kissing or groping	425	337	63.1	69.7	31.8	26.4
SV12: Flashing/unwanted exposure	117	91	73.5	82.4	18.8	16.5
SV13: Made you show your body parts	37	30	73.0 †	20.0	18.9 †	70.0
SV14: Unwanted sexual pictures or movies	81	20	74.1	65.0	17.3	30.0

Table 9-2. Proportion of unique in scope incidents reported in victimization screener based on screener items, by mode of interview for females ages 18-49 in the volunteer sample, 2014-2015 (continued)

Sexual victimization screener: Item number and description^e	Out of scope^c		Missing date information^d	
	ACASI^f	CATI^{g*}	ACASI^f	CATI^{g*}
Rate across all screening items ^h	5.8 %†	3.1 %	0.4 %	0.3 %
Rape ⁱ	4.4 %	3.4 %	0.3 %!	--
SV 1: Forced vaginal sex	4.2	4.1 !	1.0 !	--
SV 2: Forced oral sex	1.3 !	4.2 !	--	--
SV 3: Forced anal sex	--	--	--	--
SV 4: Forced digital penetration	6.1 !	6.9 !	--	--
SV 5: Penetration while unable to consent	6.3	1.6 !	--	--
Other unwanted sex ^j	10.1 %†	2.7 %	0.2 %!	--
SV 6: Sex by coercion	13.1	4.3 !	--	--
SV 7: Other unwanted sex	13.0 !	3.8 !	--	--
SV 8: Attempted unwanted sex	9.5 †	2.0 !	0.2 !	--
Sexual contact ^k	4.8 %†	3.1 %	0.6 %	0.5 %
SV 9: Unwanted kissing	5.8	2.7	--	1.1 !
SV10: Unwanted groping	3.9	2.8	0.5 !	0.5 !
SV11: Attempted unwanted kissing or groping	4.0	3.6	1.2	0.3 !
SV12: Flashing/unwanted exposure	6.8 †	1.1 !	0.9 !	--
SV13: Made you show your body parts	8.1 !	10.0 !	--	--
SV14: Unwanted sexual pictures or movies	8.6	5.0 !	--	--

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

^aIncludes incidents that happened within the 12 months prior to the interview and were eligible for a detailed incident form.

^bIncludes incidents that the respondent indicated happened as part of another incident already mentioned.

^cIncludes incidents that happened outside of the 12-month reference period.

^dIncludes incidents for which respondents could not recall the date and could not confirm whether it happened during the 12-month reference period.

^eRefers to the screening item number and the content of the item for each sexual victimization screener item. See Appendix B for full question wording.

^fAudio computer-assisted self-interview.

^gComputer-assisted telephone interview.

^hCombined total across all screening items.

ⁱIncludes screening items measuring penetrative behaviors using force.

^jIncludes screening items measuring penetration using coercion, attempted penetration, and other unwanted penetration.

^kIncludes non-penetrative behaviors using any tactic.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Similar to the pattern in the general population sample, the overall proportion of incidents that were determined to be duplicates of previously reported incidents was about 20 percent (19.5% for ACASI, 23.3% for CATI), but unlike the general population sample, the difference between modes is statistically significant. However, the magnitude of the difference is not large (19.5% ACASI vs. 23.3% CATI). The largest differences are for other unwanted penetration, where 21 percent of the CATI reports were duplicative with a prior incident compared to 9.1 percent for the ACASI.

Also, just as with the GP sample, ACASI respondents were more likely to date incidents as occurring outside the reference period when compared to CATI respondents (5.8% vs. 3.1%). Many of the individual differences are not statistically significant. One notable difference is attempted unwanted sex (9.5% ACASI vs. 2.0% for CATI).

9.1.2 Incident Date Distribution and Telescoping

Retrospective surveys, such as the RSA Pilot Test, as well as most other victimization surveys, ask respondents to remember and place events. These are, in fact, difficult cognitive tasks (Tourangeau et al., 2000). Two common types of error made by respondents are omissions and telescoping. Omissions occur for several reasons, including the respondent failing to find the event in memory (forgetting), the respondent searches memory for the wrong type of memory (mismatch), or the respondent intentionally does not report it (motivated misreporting). The RSA Pilot Test was designed to address each of these three sources of error. The use of multiple screening items, with specific descriptions, was intended to reduce forgetting by triggering different types of memories. The specific descriptions were also intended to minimize mismatches between the survey definition and the definitions used by respondents. To minimize motivated misreporting, the procedures tried to maximize confidentiality (self-administered; anonymity; assurances of confidentiality, including protection from external requests). The ACASI version of the survey also asked respondents to fill out an event history calendar, which was intended to help the respondent remember and date events in the reference period.

Telescoping is the extent to which the respondent misdates when the event occurred. Precisely dating an event is a difficult task. Dates or time periods are not typically used when storing an event in memory. One way events are placed in time is association with landmark events that have personal meaning to the respondent. Examples might be the birth of a child, moving residences or the death of a family member. Landmark events can be used to date other events. The

event history calendar was intended to assist the respondent by anchoring incidents around landmark events. Internal telescoping refers to erroneously dating an event but keeping it within the reference period covered by the survey. An example for the RSA Pilot Test is if the respondent reported an event as occurring 6 months prior to the interview when it actually occurred 2 months before. External telescoping is when respondents remember an event that occurred outside the reference period and misdate it as occurring in the reference period. For victimization surveys, this can be a significant source of error (Cantor, 1989). It is more likely to occur for salient events, which are more likely to be remembered in the first place.

To assess how respondents remembered and dated incidents, incidents were plotted by the month of the reference period. If memory was perfect, the distribution by month should reflect the actual victimization rates by month. Recall error related to omissions and telescoping can distort this distribution. With respect to omissions, one would expect that the percentage of victimizations reported by month to decline when moving in time away from the time of the interview. This would also occur if respondents misdate events by bringing them closer to the interview. External telescoping tends to be reflected by the opposite pattern. A disproportionate number of events are reported toward the beginning of the reference period.

If there is a landmark event marking the beginning of the reference period, external telescoping is less likely. For example, in longitudinal surveys, a prior interview serves as a distinctive landmark event that minimizes this type of error. The event history calendar used for the ACASI was intended to assist respondents not only to remember but also to date events. The CATI interview used an internal bounding procedure by first asking about events over the lifetime and then asked about events that occurred in the last 12 months. This method has also been found to reduce external telescoping (Sudman, et al, 1984; Loftus et al, 1990).

In this section, the distribution by month is analyzed by the type of victimization reported and mode of interview. The results are discussed with respect to their implications for telescoping for each mode of interview.

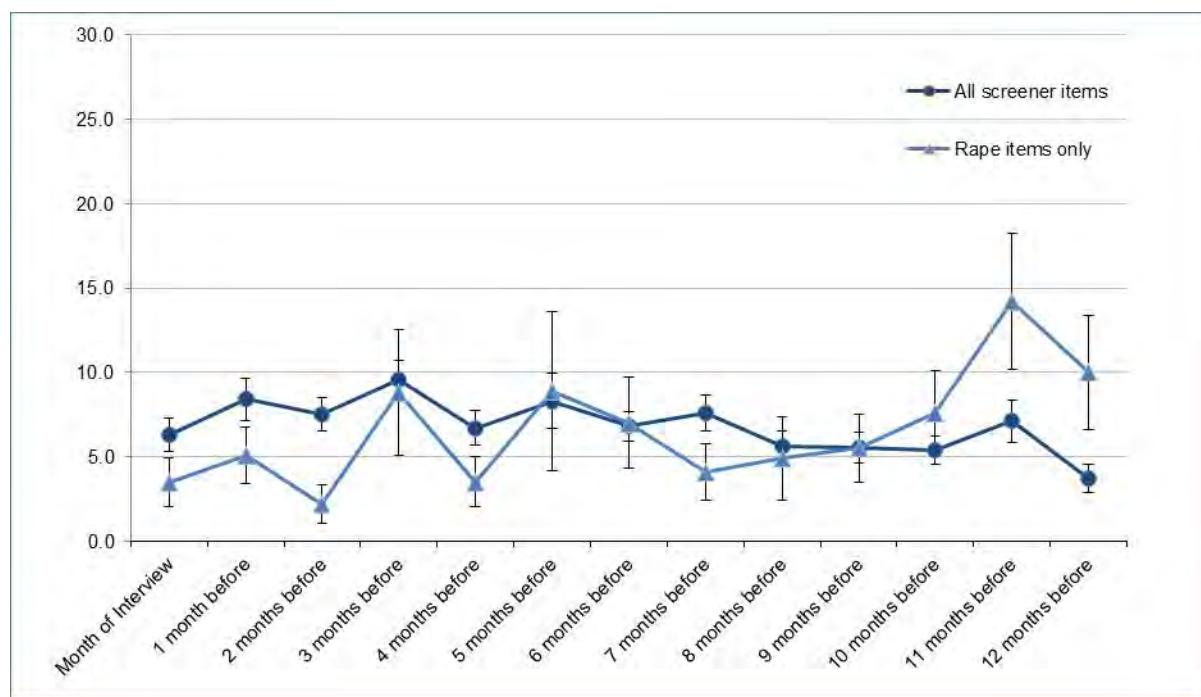
General Population

Figures 9-1 and 9-2 provide the percent of incidents reported by the month of the reference period for the general population sample for each mode. The horizontal axis is the month of the reference period, with 0 being the month of the interview and 12 being the beginning of the period.

One line provides the percent of all the incidents reported in each month. The second line is for incidents reported for the completed rape screening questions (SV1 – SV5). Overall, both modes showed a slight decline in incidents over the course of the reference period for the line representing all incidents, with slightly more incidents being reported in the most recent months than in the more distant months. This is consistent with respondents forgetting incidents that occurred when moving from the most recent to the most distant time periods. It may also be indicative of some forward telescoping, which involves respondents misdating events as occurring later than actually occurring.

The second line representing rapes is somewhat different across the two modes. For ACASI, the two highest percentages are for the two earliest months of the reference period (months 11 and 12). These 2 months represent almost one in four of the incidents. This is consistent with external telescoping for these more serious types of incidents.

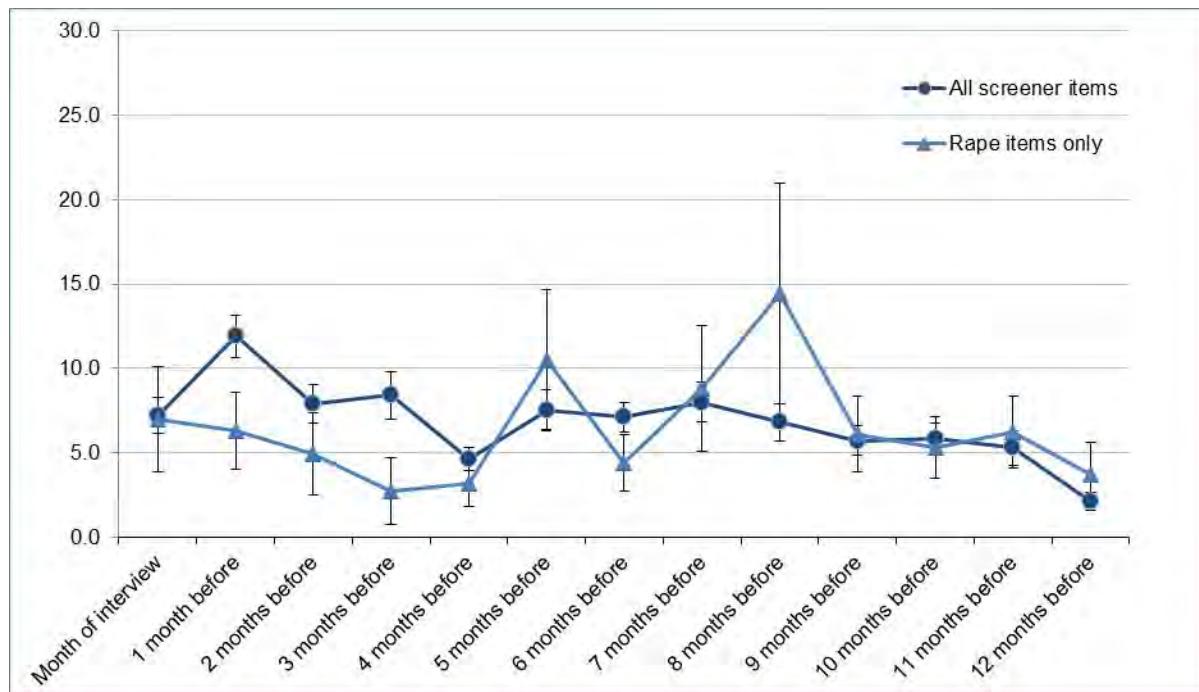
Figure 9-1. Recency distribution of incident dates reported in the ACASI interview mode by type of victimization for females age 18-49 in the general population



Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

The pattern among CATI respondents for rape incidents is less clear. Although these incidents do not show the same forgetting curve as less serious incident types, they also do not clearly show external telescoping effects, with only 3.9 percent of respondents saying that the incident occurred in the earliest 2 months of the reference period. The two most common dates given for CATI incidents of this type were 5 months and 8 months prior to the interview.

Figure 9-2. Recency distribution of incident dates reported in the CATI interview mode by type of victimization for females age 18-49 in the general population



Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

Interpretation across modes is complicated by the different calendar periods when data collection occurred among the two general population samples. The vast majority of the ACASI was collected from November of 2014 to June 2015,⁴² while the vast majority of the CATI was collected from June to September of 2015.⁴³ This means that early and later months of the reference periods for the ACASI modes are somewhat different than for the CATI. To the extent there is a seasonal fluctuation in the rate of sexual assault, comparing the patterns by month across modes is not entirely straightforward. Analysis of NCVS data, which has a fully balanced sample by month of interview, does not display strong seasonal variation for rape and sexual assault, however (Lauritsen & White, 2014). The seasonality that does exist for slightly higher rates of rape and sexual assault in the summer months (June, July, and August). Across the two modes, the CATI had a higher percentage of interviews in which the early months of the reference period were in the summer, which is the opposite of what is shown in the charts. The ACASI displayed a higher percentage of rapes in these months, which bolsters the interpretation of external telescoping for the ACASI.

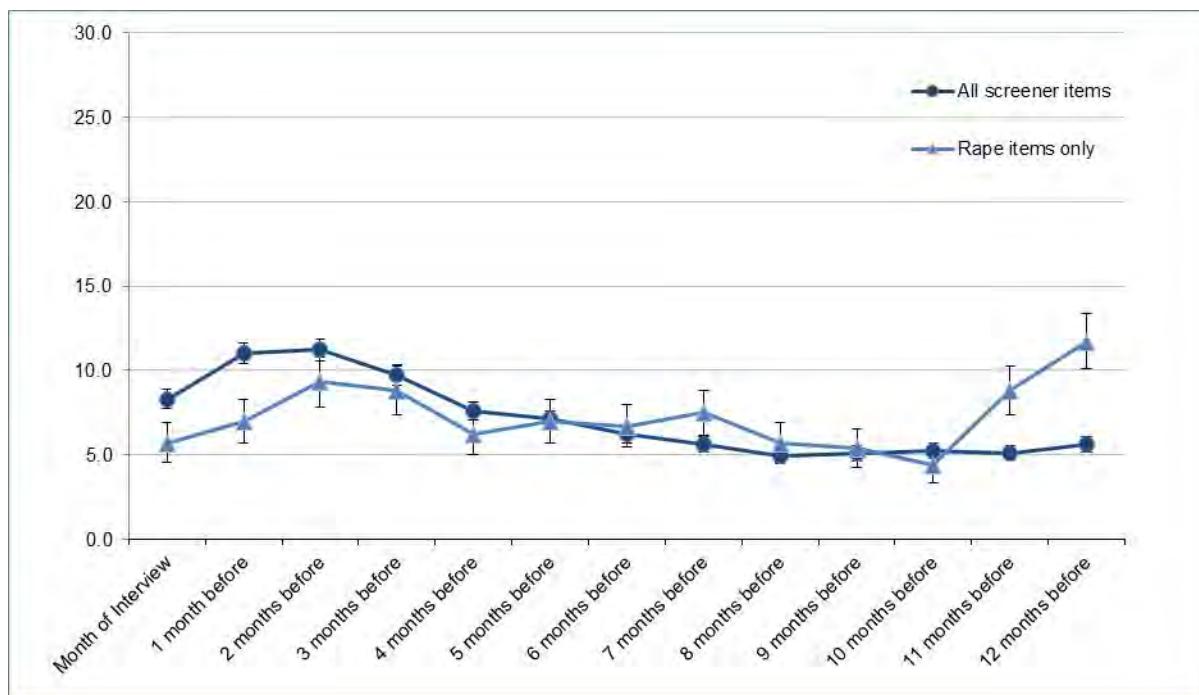
⁴²A small percentage of data was collected in October of 2014.

⁴³A small percentage of data was collected in May and October of 2015.

Volunteer Sample

The timing of the interviews for the volunteer samples was closer between the two modes. For the ACASI, the interviews took place between October of 2014 and January of 2015. The ACASI was similar with 80 percent of the sample being interviewed in this time period. The remaining 20 percent was interviewed between February and May of 2015. This significantly reduces any concerns about the role that seasonality might account for in creating any differences between the two modes. The results are similar to those in the general population (figures 9-3 and 9-4).

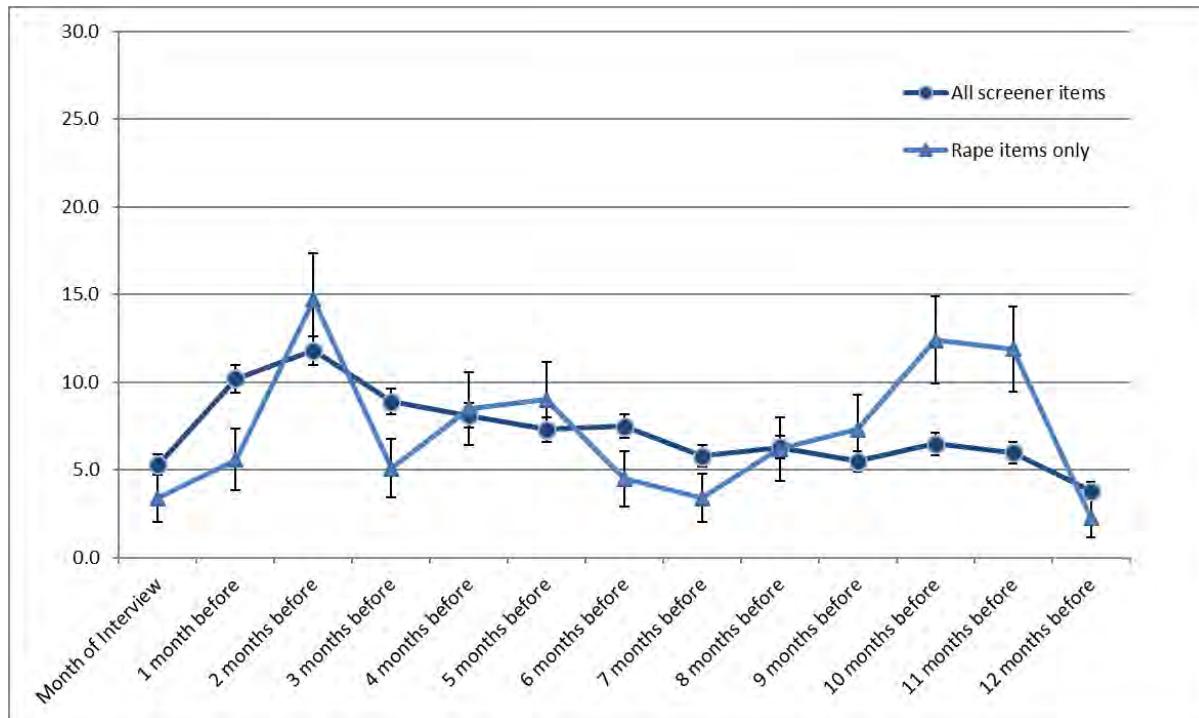
Figure 9-3. Recency distribution of incident dates reported in the ACASI Interview mode by type of victimization for females age 18-29 in the volunteer sample



Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

For all incidents, there is a slightly higher proportion reported in the most recent months than in more distant months. However, also similar to general population patterns, rape incidents do not follow this general trend. Respondents in the ACASI condition show evidence of telescoping when reporting the months in which rape incidents occurred, with 20.5 percent of incidents reported in the first 2 months of the reference period (figure 9-4).

Figure 9-4. Recency distribution of incident dates reported in the CATI interview mode by type of victimization for females age 18-29 in the volunteer sample



Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

The patterns in CATI results are less clear with very wide variation in the proportion reported in each month. There is a relatively large percentage of incidents in months 10 and 11 of the reference period (24.3%). However, there is a significant drop in month 12, which has one of the lowest proportions. There is also a significant proportion reported in months 2 and 3.

Summary

In summary, there is evidence of external telescoping for the completed rape estimates for the ACASI. For both the general population and the volunteer samples, the highest proportion of incidents were reported 11 and 12 months from the interview. As noted above, it is also the case that for the ACASI interviews a higher proportion of initial reports of incidents were dated as outside the reference period, which can also be interpreted as an indication of external telescoping when answering the SV screener items. For the CATI interviews, the evidence was mixed. For the general population there were no indications of external telescoping. The most frequently reported months were in the middle, rather than the beginning, of the reference period. For the VO sample,

months 10 and 11 were high, indicating some telescoping. However, month 12 was very low, which is inconsistent with external telescoping.

At least from this evidence, the event history calendar used for the ACASI survey was not effective at reducing external telescoping. The CATI procedure, which used the lifetime question to serve as an internal bound for the 12-month question, seemed to be more effective at minimizing this type of error.

9.1.3 Missing Data

This section describes the extent of missing data on the SV screener.

SV Item Nonresponse

The rate of missing data due to item nonresponse in the SV screener was very low. Across both modes, all SV screener items had far less than 1 percent of general population respondents saying “don’t know” or declining to give a response. The screener item with the highest rate of missing data was for ACASI item SV8 (attempted unwanted penetration – 0.5% missing). ACASI respondents had a significantly higher rate of item missing data for 3 out of the 14 screener items (see Appendix G for nonresponse rates to individual SV items), but differences in the screener overall were not significant. The follow-up items in the SV screener, which asked for the month in which the incident occurred, had relatively low rates of missing data. Less than 2 percent of incidents reported by the general population screener were missing date information.

Rates of item nonresponse among the volunteer sample were even lower than what was seen in the general population sample, with an overall missing rate of 0.1 percent for both ACASI and CATI. To see nonresponse rates for all SV items among the volunteer sample, see Appendix G.

No Date Due to Four-Incident Cap

If a respondent said an incident occurred in the last 12 months, she was asked how many times it occurred. A month and year was collected for up to four incidents. An incident was considered eligible for a DIF if it was specifically dated as occurring with the last 12 months. If the respondent reported more than four incidents, the survey asked if the remaining items occurred

within the last 12 months. However, these items were not considered for a DIF. This capping procedure was used to minimize burden on the respondent. Nevertheless, it does omit some incidents from rate calculation based on the DIF.

Overall, the four-incident cap on follow-up questions to each SV item was sufficient for about 85 to 90 percent of general population respondents who reported victimization. This rate was highest for the items asking about rape, where about 90 to 95 percent of respondents who reported these types of victimization were able to provide all of the incident dates that they remembered within the past year (91.2% for ACASI, 94.8% for CATT) (table 9-3).

Table 9-3. Percent of respondents saying yes to each screener item who indicated the incident happened more than four times in the past 12 months by screener item and mode of interview for females ages 18-49 in the general population, 2014-2015

Sexual victimization screener: Item number and description ^a	Number of weighted sample cases			
	ACASI ^b	CATI ^c	ACASI ^b	CATI ^{c*}
Rate across all screener items ^d	1,609,022	1,294,811	10.3%	15.6%
Rape ^e	359,734	264,550	8.8%!	5.2!
SV 1: Forced vaginal sex	127,769	78,750	8.1!	6.5!
SV 2: Forced oral sex	95,037	52,228	--	6.3!
SV 3: Forced anal sex	50,248	49,027	--	--
SV 4: Forced digital penetration	97,630	29,273	21.6!	13.4!
SV 5: Penetration while unable to consent	176,037	144,340	--	3.4!
Other unwanted sex ^f	561,152	325,244	6%	10.6%
SV 6: Sex by coercion	48,824	47,715	28.1!	10.7!
SV 7: Other unwanted sex	72,718	31,586	0.5!	12.4!
SV 8: Attempted unwanted sex	500,249	273,145	4	9.3
Sexual contact ^g	1,348,372	1,160,362	8.9%	14.4%
SV 9: Unwanted kissing	328,807	285,994	3.3!	4.5!
SV10: Unwanted groping	799,373	637,098	10.4	14.6
SV11: Attempted unwanted kissing or groping	661,474	627,958	5.4!	9.9
SV12: Flashing/unwanted exposure	196,868	183,070	4.8!	3.7!
SV13: Made you show your body parts	50,006	54,788	7.6!	6.4!
SV14: Unwanted sexual pictures or movies	73,113	49,091	9.8!	10.1!

Note: Estimates are based on weighted data. See Appendix A for standard errors.

*Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

^aRefers to the screening item number and the content of the item for each sexual victimization screener item. See Appendix B for full question wording.

^bAudio computer-assisted self-interview.

^cComputer-assisted telephone interview.

^dCombined total across all screening items.

^eIncludes screening items measuring penetrative behaviors using force or while unable to consent due to alcohol or drugs.

^fIncludes screening items measuring penetration using coercion, attempted penetration, and other unwanted penetration.

^gIncludes non-penetrative behaviors using any tactic.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

This percentage was similar, although slightly lower among respondents reporting victimization in the volunteer sample, where about 84 percent of respondents were given the opportunity to report a date for all incidents experienced in the past year (see Appendix G).

Although the vast majority of respondents provided a date for all incidents, there were a significant number who did not. For example, 8.8 percent of those reporting to the completed rape items said the incident occurred more than four times. While this is sufficient for estimating prevalence, not counting these incidents will lead to an underestimate of the incidence of these victimizations. We do not recommend collecting dates on more incidents. Rather, the survey should consider some type of imputation procedure to estimate the impact of this capping on the estimates (chapter 15).

9.2 Detailed Incident Form Results

9.2.1 DIF Eligibility and Completion

In the RSA Pilot Test, respondents were asked about 14 types of unwanted sexual contact. Incidents that occurred within the past 12 months were eligible for a detailed incident form (DIF), which was used to collect information on the characteristics of incidents. A maximum of three DIFs were collected for each respondent. This section provides data on the extent to which respondents filled out the requested DIFs.

If the respondent reported more than three incidents on the screener, the instrument was programmed to prioritize incidents. The highest priority was for penetration by force or inability to consent, followed by other unwanted penetration, unwanted sexual touching and unwanted sexual non-contact (e.g., exposure). If a respondent declined to fill out the details for a particular incident, or did not complete the DIF, the program moved to the incident that was next on the priority list. This was done for up to five incidents until three incidents were completed.

General Population

For the analysis in this section, a DIF is considered complete if enough questions have been answered to assign a classification category or the respondent was administered every item used in the classification process. Overall, the majority of the general population completed all of the DIFs that they were eligible for on ACASI (table 9-4). Specifically, 99.3 percent of ACASI general population respondents who were eligible for one DIF, 97.8 percent of respondents who were eligible for two DIFs, and 93.2 percent of respondents who were eligible for three DIFs completed

all of their eligible DIFs. Among those who reported four eligible DIFs, 87.4 percent of respondents completed three DIFs. Most of those that did not fill out all three DIFs only completed one. For those who had five or more eligible DIFs, 81.7 percent completed three DIFs. For those who did not do three, most filled out two. There were very few individuals who did not complete any DIFs.

Table 9-4. Number of DIF incidents completed by number of incidents eligible for a DIF and mode of interview for females ages 18-49 in the general population, 2014-2015

Number of incidents eligible for DIF ^d	Number of weighted sample cases		ACASI ^b			
	ACASI ^b	CATI ^c	0	1	2	3
1	670,857	432,141	0.7 %!	99.3 %	-- %	-- %
2	228,240	227,297	--	2.2 †!	97.8 †	--
3	188,257	124,499	--	2.1 !	4.7 !	93.2 †
4	136,438	101,588	--	10.9	1.7 !	87.4
5 or more	256,513	265,715	1.2 !	3.6 †!	13.5	81.7 †

Table 9-4 (continued)

Number of incidents eligible for DIF ^d	Number of weighted sample cases		CATI ^{c*}			
	ACASI ^b	CATI ^c	0	1	2	3
1	670,857	432,141	1.9 %!	98.1 %	-- %	-- %
2	228,240	227,297	6.1 !	11.1	82.7	--
3	188,257	124,499	1.8 !	16.7	16.5	64.9
4	136,438	101,588	5.5 !	18.4	13.9 !	62.2
5 or more	256,513	265,715	4.6 !	20.7	14.8	60.0

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

† Significant difference from CATI at the 95% confidence level.

^aIncludes DIFs where respondent was either administered all classification questions or gave enough information to classify the incident.

^bAudio computer-assisted self-interview.

^cComputer-assisted telephone interview.

^dIncludes the number of detailed incident forms the respondent was eligible to respond to, based on responses to the sexual victimization screening items.

^eIncludes the number of detailed incident forms actually completed by the respondent, regardless of how many they were eligible for.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

For CATI, 98.1 percent of respondents who had one eligible DIF filled one out. After that, there is a significant decline in the number of respondents who filled out all of the requested DIFs. Of those with two eligible DIFs, 82.7 percent filled out both. Of those with three eligible DIFs, 64.9 percent filled out all three. Among those with four or five eligible DIFs, 62.2 percent and 60.0 percent, respectively, filled out three.

There is a mode difference between completion rates. The general population DIF completion rates on CATI are significantly lower than the ones on ACASI for two (82.7% CATI vs. 97.8% ACASI), three (64.9% CATI vs. 93.2% ACASI), and five (60% CATI vs. 81.7% ACASI) eligible DIFs. Although not statistically significant, there is a higher percentage of respondents who are eligible for at least one DIF but completed none in CATI compared to ACASI. For example, 3.8 percent of the CATI respondents did not provide any DIFs (12/311), which compares to only 0.4 percent of the ACASI respondents.

Volunteer Sample

Table 9-5 shows the distribution of completed by eligible DIFs for the volunteer sample. Overall, the DIF completion rates for one or two eligible DIFs are relatively high. For ACASI, 100 percent of respondents who were eligible for one DIF completed one DIF and 97.4 percent who were eligible for two DIFs completed two DIFs. It drops somewhat for those who are eligible for three DIFs to 79.7 percent. Most of this dropout occurs when going from two to three DIFs. Among those who were eligible for four or more DIFs, the percentage that completed all three DIFs goes back up (88.9% and 95.1%).

The pattern for CATI is similar to what is described above for the GP sample with fewer respondents filling out all three DIFs when compared to the VO ACASI. Practically all respondents (99.0%) who were eligible for one DIF completed one DIF. It goes down to 87.7 percent for those who were eligible for two DIFs and completed two DIFs. The DIF completion rate for those who were eligible for three DIFs goes down to 70.8 percent. The percent of those who have four or more eligible incidents is similar to those with three eligible incidents (68.4% and 78.0%). Also similar to the general population, CATI has more respondents who were eligible for DIFs but completed none when compared to ACASI.

Table 9-5. Number of DIF incidents completed by number of incidents eligible for a DIF and mode of interview for females ages 18-29 in the volunteer sample, 2014–2015

Number of incidents eligible for DIF ^d	Number of unweighted sample cases		ACASI ^b			
	ACASI ^b	CATI ^c	0	1	2	3 ^f
1	169	105	-- %	100.0 %	-- %	-- %
2	76	65	-- !	2.6 !	97.4 †	--
3	59	48	-- !	5.1 !	15.3	79.7
4	54	38	-- !	3.7 !	7.4	88.9 †
5 or more	163	82	--	2.5	2.5 †	95.1 †

Table 9-5 (continued)

Number of incidents eligible for DIF ^d	Number of unweighted sample cases		CATI ^{c*}			
	ACASI ^b	CATI ^c	0	1	2	3 ^f
1	169	105	1.0 %!	99.0 %	-- %	-- %
2	76	65	3.1 !	9.2	87.7	--
3	59	48	4.2 !	8.3	16.7	70.8
4	54	38	2.6 !	13.2	15.8	68.4
5 or more	163	82	2.4 !	7.3	12.2	78.0

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

-- Less than 0.05%.

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

^aIncludes DIFs where respondent was either administered all classification questions or gave enough information to classify the incident.

^bAudio computer-assisted self-interview.

^cComputer-assisted telephone interview.

^dIncludes the number of detailed incident forms the respondent was eligible to respond to, based on responses to the sexual victimization screening items.

^eIncludes the number of detailed incident forms actually completed by the respondent, regardless of how many they were eligible for.

^fIncludes three respondents who completed four DIFs.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Discussion of Completion of the DIFs

The amount of missing data due to respondents not completing a DIF was different across the two modes. For the ACASI, about 90 percent of the respondents filled out all of the DIFs that were requested. When including those who filled out three DIFs, but had to move to the fourth or fifth eligible DIF, this percentage goes up to 94 percent. For the CATI, these percentages were 67 and 80 percent, respectively. The higher rate of missing DIFs for the CATI is likely due to the amount of time to complete this portion of the survey. As noted in the section on the survey timings (section 6.1), the CATI respondents took much longer to fill out a DIF; one reason is that the interviewers were required to read out not only the questions, but also the response categories.

9.2.2 Missing Data

This section discusses item nonresponse in the detailed incident form (DIF) for the general population and volunteer sample by mode and detailed incident number. For item nonresponse, each individual question in the DIF was reviewed by looking at the ratio of respondents providing a missing response (i.e., “don’t know” or refused) to the number of eligible respondents. A full list of the item nonresponse rates can be found in Appendix G. In this section, we focus on the patterns for the general population sample. The patterns for the volunteer sample were similar. This section also describes the results of a feature that was built into the DIF. If respondents provided 10 or more “don’t know” or refused answers on the DIF or answered all of the incident description questions (D1a-D1d and D2a-D2e) with “don’t know” or refused, they were skipped to the end of the DIF and the incident narrative was collected before moving to the next part of the instrument.

Overall, for most of the items on the DIF the item nonresponse was very low. For ACASI respondents, 87 percent of the questions had less than 5 percent missing data (i.e., 163 DIF questions had less than 5 percent missing data out of a total of 187 DIF items). This compares to 94.7 percent for CATI respondents (i.e., 177 questions had less than 5 percent missing data out of a total of 187 DIF items).

For ACASI respondents, 50 percent of the questions with an item nonresponse rate of 5 percent or more were open-ended “other specify” questions (12 of the 24 items with a nonresponse rate of 5 percent or more). For example, 29.1 percent was missing for a question about what other types of unwanted sexual contact that may have occurred (Appendix G). Similarly, the rate was 42.1 percent for a question asking about other types of relationships the respondent had with the

offender (Appendix G). The most missing data for closed-ended questions are in section D. The questions on unwanted sexual behavior (D1a – d; D2a – d) had missing rates that were between 6 percent and 8 percent. The question on the type of physical force used (D8) had a rate of 13.9 percent. Outside of Section D, the question on whether the offender was using alcohol or drugs also had higher missing data (6.9%).

Among CATI respondents, 20 percent of the items with a nonresponse rate of 5 percent or more were open-ended (i.e., 2 of the 10 items with a nonresponse rate of 5 percent or more). The rest were closed-ended questions that had missing rates between 5 percent and 10 percent. As with the ACASI, respondents were more likely to not answer the question on the offender's use of alcohol or drugs (15.5%). Unlike the ACASI, the items on unwanted behavior (D1, D2) did not have high rates of missing data, nor did the item on the type of physical injury (D8).

Item nonresponse generally increased as respondents were administered additional DIFs. For example, among ACASI respondents who were asked about what they were doing at the time of the incident (C7), there was less than 1 percent nonresponse on the first DIF administered, 4.9 percent on the second DIF administered, 18.7 percent on the third DIF administered, and 86.9 percent on the fourth DIF administered (Appendix G). For CATI respondents on the same question, the nonresponse rate was 1.5 percent on the first DIF, 3.1 percent on the second DIF, and 15.2 percent on the third DIF.

As noted above, if a respondent skipped 10 or more questions or did not answer any of the behavior questions (D1, D2), she was just asked to describe the incident in a narrative. Overall, 8.1 percent of DIFs filled out on ACASI and 3.1 on CATI triggered this nonresponse threshold. Among the ACASI respondents who triggered the threshold, 29.8 percent provided enough information to classify the incident into a type of crime category. The percentage providing enough information on the CATI was smaller (9.7%).

9.3 Comparing BSQ, DIF and Narratives

This section compares three different measures of sexual assault from the survey – the behaviorally specific screening items (BSQs), the information from the detailed incident form (DIF), and the detailed narratives that were requested at the end of each DIF. These comparisons were done for two reasons. First, the review provides a qualitative idea of the types of incidents that were

reported on the RSA Pilot Test. As noted in the previous chapter, the incidence rates of rape for the RSA Pilot Test are 60 to 70 times higher than the NCVS. This raises some questions about the types of incidents that are being reported on the Pilot. The RSA Pilot Test measures a wider range of incidents than NCVS (e.g., alcohol/drug facilitated), but this alone does not account for the huge differences in the rates. By providing the narrative descriptions of the incidents below, it is possible to get a concrete idea of what was reported and how it relates to legal definitions of rape.

A second purpose of this section is to cross-check measures that should be measuring the same set of behaviors and tactics. The RSA Pilot Test used a two-stage design. The first stage administered 14 BSQs covering different types of sexual violence. The second stage administered the DIF, which contains descriptive information about the incident. BSQs are relatively complex, as they specify the behavior, the tactic, and the extent to which consent was given in a single question. Some respondents have difficulty encoding and processing all of these elements when answering (Steiger et al, 2014; Cantor et al, 2014; Cook et al., 2011). The DIF permits separate collection of each element needed to classify and describe the event. However, the BSQ provides a summary that may describe the respondent's overall experience in way that is easier to understand than parsing the elements of the incident into its component parts.

The narrative provides a third measure of what happened during the incident. If a complete narrative is provided, it can be used as a check on what was collected on the BSQ and the DIF. For example, the NCVS edits the DIF using the narrative provided at the end of the incident form. In some cases, the narrative reveals problems with the measure on the DIF. Respondents may have misunderstood certain questions (e.g., did it occur at home?), which subsequently leads to misclassification because relevant questions are not answered correctly.

By comparing the BSQ, DIF, and narrative, the analysis below provides information on possible errors at each of the two stages of collection. This information is used as one method to assess the quality of the data that was collected at each stage and to make recommendations for future surveys on RSA.

The SV screener consists of 14 BSQs that cover a range of nonconsensual sexual behaviors from forced vaginal penetration to exposure. Questions in the SV were asked using slightly different approaches between modes. For each of the SV screener items, respondents in the CATI condition were first asked whether they had ever experienced an incident of that type in their lifetime. Those who said "yes" were asked, "Has this happened at any time since [month/year]?" where "month/year" is the month of the interview for the previous year.

By contrast, respondents in the ACASI condition were first asked whether they had experienced a sexual victimization in the past 12 months. If the answer was “yes,” they were asked which month and year it occurred. To date the incident, the ACASI respondents were provided a list of 14 month/year combinations. Thirteen enumerated each month of the reference period (including the month of the interview). The last category was for the time period before the beginning of the reference period. If the respondent chose this option, the incident was not counted as occurring in the reference period. At the end of the SV screener, the ACASI respondents were asked about a lifetime victimization if they had said “no” to previous past 12-month victimization of that type.

Table 9-6 shows the wording of each of the SV screener items. This table shows the ACASI version of the items, which refers to past 12-month victimizations. CATI items differ only in that they ask if respondents have ever experienced that type of victimization. The last column provides the type of incident the screener was intended to capture. The wording of the SV screener for incidents that did not involve rape (i.e., completed penetration by force or while unable to consent) includes incidents that refer to a wider range of unwanted incidents, although all but SV6 (sex by coercion) ask if the incident occurred “against your will.” As noted in Chapter 2, the SV screener was not intended to be used as the final classification of the incident. Nonetheless, generating rates by these items is instructive when comparing to the results from the DIF, especially for the estimates of completed rape, where the intent of the SV screener and DIF do correspond.

Table 9-6. Sexual victimization screener question wording for ACASI

Item	ACASI question wording	Type of Incident
SV1	Since [anchordate], has a male used force or threats of force to make you have vaginal sex against your will? By vaginal sex, it means putting his penis in your vagina.	Completed forced rape (vaginal)
SV2	Since [anchordate], has anyone, male or female, used force or threats of force to make you have oral sex against your will? By oral sex, it means that someone penetrated your vagina or anus with their mouth or tongue, or you were forced to use your mouth or tongue on someone else's genitals or anus.	Completed forced rape (oral)
SV3	Since [anchordate], has a male used force or threats of force to make you have anal sex against your will? By anal sex, it means that a man or boy put his penis in your anus.	Completed forced rape (anal)
SV4	Since [anchordate], has a male or female used force or threats of force to put fingers or a foreign object in your vagina or anus against your will?	Completed forced rape (digital)
SV5	(Other than the incidents you have already mentioned), since [anchordate], has anyone made you have any type of sex when you were unable to consent because you were too drunk, high, or passed out?	Completed rape while unable to consent
SV6	(Other than the incidents you have already mentioned), since [anchordate], has anyone made you have any type of sex by threatening to cause problems for you, such as at your job or school, at home, in your relationships, or in any other way?	Coercive penetration
SV7	(Other than the incidents you have already mentioned), since [anchordate], have you been in <u>any other situations</u> where someone made you have any type of sex against your will?	Other penetration against will
SV8	Thinking about all the different types of situations you have been asked about so far, since [anchordate], has anyone <u>tried, but did not succeed</u> , at making you have any type of sex against your will?	Attempted penetration against will
SV9	(Other than the incidents you have already mentioned), since [anchordate], has anyone, male or female, kissed you in a sexual way against your will?	Kissing against will
SV10	(Other than the incidents you have already mentioned), since [anchordate], has anyone, male or female, fondled, groped, grabbed, or touched you against your will?	Other sexual contact against will
SV11	Since [anchordate], has anyone tried, but did not succeed at kissing, fondling, groping, grabbing, or touching you against your will?	Attempted other sexual contact against will
SV12	(Other than the incidents you have already mentioned), since [anchordate], has anyone, male or female, made you watch against your will while they exposed their sexual body parts to you, flashed you, or masturbated in front of you?	Flashing and exposure
SV13	(Other than the incidents you have already mentioned), since [anchordate], has anyone, male or female, made you show your sexual body parts to them against your will?	Show your body against will
SV14	(Other than the incidents you have already mentioned), since [anchordate], has anyone, male or female, made you look at or participate in sexual photos or movies against your will?	Making sexual videos/pictures against will

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

For the RSA Pilot Test, the BSQs can be linked to both completed rape (questions SV1 – SV5) and attempted penetration that is unwanted (SV8). The DIF can be parsed using multiple classification schemes, including incidents classified as completed rape, as well as unwanted attempted penetration. The analysis below compares the BSQ and DIF classifications of these two types of incidents and uses the narratives to provide a description of what was actually reported by the respondent.

Theoretically, there should be a link between the SV screener items and the final DIF classification. As illustrated in Table 9-7, in practice there is a discrepancy in how the two correspond. The first five items, SV1 to SV5, targeted completed rapes related to the two tactics of physical force or inability to consent. Item SV8 targeted attempted penetration against the victim's will, but did not specify the tactic. Among these 6 items, 83 percent of the incidents were classified as rapes on the DIF ($54\% + 29\% = 83\%$). The remaining incidents classified as a rape by the DIF were reported when answering one of the other 8 BSQs on the screener. The discrepancy is larger for sexual assaults. There were five items, in total, targeting non-penetrative sexual contact and non-contact (SV9 – SV14). Respondents reported approximately 75 percent of the incidents classified by the DIF from these items ($53\% + 22\% + 3\% = 75\%$), with the remaining 25 percent being reported in response to SV items asking about penetration.

In the remainder of this section, the reasons for the discrepancies between these two types of questions are discussed in more detail.

Table 9-7. Percent of rapes and sexual assaults reported from SV screener items targeting different behaviors and tactics

BSQ items targeting:	# of items	Percent of:	
		Rapes	Sexual assaults
Penetration against victim's will by force or unable to consent	5	54%	5%
Other penetration against victim's will	2	5%	1%
Attempted penetration against victim's will	1	29%	16%
Kissing, sexual touching against victim's will	2	7%	53%
Attempted touching against victim's will	1	3%	22%
Exposure, photo or videos against victim's will	3	2%	3%
Total	14	100%	100%

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

9.3.1 Comparison of Classifications for Completed Rape

In this section the correspondence between the BSQs and the DIF, and narratives are discussed for completed rapes (SV1-SV5). Section 9.3.2 discusses reports that were classified as attempted penetration from the BSQs and the DIF.

The first five screening items ask if the person used force or threats of force to do any of the following to the woman against her will: have vaginal sex (SV1), oral sex (SV2), anal sex (SV3), or digital penetration (SV4). The fifth BSQ asks about having sex while she was unable to consent because she was too drunk, high, or passed out (SV5). The discussion below compares the responses to these five BSQs intended to measure completed rape to the classification from the DIF.

Table 9-8 provides the prevalence rates from the BSQs versus the classifications based on the DIF. Using the BSQs, the past 12-month rate is 2.8 percent in ACASI and 2.2 percent in CATI. Alternately, the past 12-month prevalence of completed rape using the DIF is 2.4 percent in ACASI and 2.1 percent in CATI. While the differences in prevalence rates of completed rape between the BSQ and DIF are not significantly different, further analysis shows that there are a number of discrepancies between the two measurement approaches.

Table 9-8. Prevalence rate of completed rape by type of crime classification, mode of interview and classification method for females ages 18-49 in the general population, 2014-2015

Classification	ACASI ^a		CATI ^{b*}	
	BSQ ^c	DIF ^{d*}	BSQ ^c	DIF ^{d*}
Completed rape ^e	2.8 %	2.4 %	2.2 %	2.1 %
Forced ^f	1.7	1.9	1.3	1.7
Unable to consent ^g	1.2 †	0.5	1.1 †	0.5

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cBehaviorally specific questions used in the sexual victimization screener questions.

^dDetailed incident form questions used to classify the incident.

^eIncludes penetrative sexual contact where the offender completed the behavior using force or while unable to consent.

^fIncludes holding or pinning, using a weapon or threatening to use a weapon, other physical attacks or threats of physical attacks on respondent or someone else.

^gIncludes incidents where respondents were passed out for all or parts of the incident or were unable to consent due to alcohol or drugs.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Completed Rape on the BSQ

For the general population sample, fully half of ACASI incidents that screened in as completed rape were not classified by the DIF as a completed rape (table 9-9). For CATI, roughly one-quarter of the BSQ incidents did not get classified as a completed rape from the DIF. In both ACASI and CATI incidents, roughly 16 to 17 percent of incidents screening in as completed rape were ultimately classified as attempted rapes, and roughly 5 percent were ultimately classified as sexual assault.

There are significant differences between the two modes for the general population sample. The percent of reports that do not match to the DIF classification is about twice as large on ACASI when compared to CATI (50% vs. 25%). The discrepancy between the two modes is partly related to incidents that were not classified at all. The largest portion of these are ACASI incidents where the respondent did not report any type of unwanted behavior on the DIF. Section D of the DIF is where respondents are asked about different forms of unwanted behaviors. Among ACASI respondents, 9 percent of the respondents who responded affirmatively to SV1 – SV5 did not report any unwanted behavior in section D. This is much smaller for CATI, where only 1 percent did not report any unwanted behavior. The second group of incidents that were not classified are those for which the respondent did not provide enough behavior or tactic information to classify the event. This represents those individuals who skipped a critical item needed to classify an incident. This is 7.3 percent for the ACASI and 0.6 percent for the CATI.

The results for the VO sample are similar to that of the general population, although the differences between ACASI and CATI are not as large. Approximately 35 percent of the ACASI respondents who reported a completed rape on a BSQ were not classified as a completed rape using the DIF. For CATI, this is 25 percent. The ACASI respondents had a similar, but not as dramatic, pattern related to incidents that were not classified into a particular type of unwanted behavior. Five percent of the ACASI respondents did not report an unwanted behavior, while 3.6 percent did not provide answers to the items on either the behaviors or tactics. It is virtually zero among the CATI responses.

Table 9-9. Distribution of detailed incident form classifications among incidents classified by behaviorally specific questions as completed rape, by sample type and mode for females, 2014–2015

DIF ^c classification	BSQ ^a classification: Completed rape ^b (SV1-5)			
	General population		Volunteer sample	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e *
Completed rape ^c	50.8 %†	74.3 %	65.7 %	74.3 %
Forced ^f	39.0	54.9	54.3	61.9
Unable to consent ^g	11.7	19.3	11.4	12.4
Attempted and threatened rape ^h	17.3 %	16.2 %	9.9 %	13.3 %
Sexual assault ⁱ	4.6 %!	5.0 %!	3.5 %	2.7 %!
Other unwanted sexual contact ^j	11.0 %	3.0 %!	12.2 %	8.9 %
Not enough information to classify ^k	16.4 %†	1.6 %!	8.7 %†	0.9 %!
No unwanted behavior	9.0 †	1.0 !	5.1	--
Not enough behavior information to classify ^l	5.0 !	0.6 !	2.0	--
Not enough tactic information to classify ^m	2.3 !	--	1.6	0.9 !
Number of unweighted sample cases	120	63	254	113

Note: General population estimates are based on weighted data for ages 18–49. Volunteer sample estimates are based on unweighted data for ages 18–29. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%

^aBehaviorally specific questions used in the sexual victimization screener questions.

^bIncludes sexual victimization screening items about vaginal, oral, anal, and digital forced penetration and penetration while unable to consent due to alcohol or drugs.

^cDetailed incident form questions used to classify the incident.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fIncludes holding or pinning, using a weapon or threatening to use a weapon, other physical attacks or threats of physical attacks on respondent or someone else.

^gIncludes incidents where respondents were passed out for all or parts of the incident or were unable to consent due to alcohol or drugs.

^hIncludes penetrative sexual contact using force or while unable to consent where the offender verbally threatened or physically attempted, but did not complete the behavior.

ⁱIncludes non-penetrative sexual contact using force or while unable to consent.

^jIncludes unwanted penetrative or non-penetrative sexual contact where force or while unable to consent was not reported while unable to consent.

^kIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all behavior and tactic items.

^lIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all behavior items.

^mIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all tactic items.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

Completed Rape on the DIF

The comparison above examined completed rapes identified by a BSQ to the DIF classification. There were also a significant number of incidents identified as a completed rape on the DIF that were not classified as such by the BSQ. Among the incidents that were classified as completed rape by the DIF, 30 to 35 percent screened in as something other than a completed rape (table 9-10). Results were consistent across modes, with 15 to 20 percent of completed rapes screening in as some other type of unwanted sex (SV6 through SV8) and 15 percent screening in as incidents involving kissing, groping, or some other form of sexual touching or exposure (items SV9 through SV14).

Results were comparable for volunteer (VO) respondents, with 70.2 percent of completed rapes having screened in from items SV1 through SV5 in the ACASI mode. CATI volunteers were slightly more consistent between the SV screening items and the DIF classification, with 77.0 percent of completed rapes screening in from items SV1 through SV5.

Table 9-10. Distribution of behaviorally specific question classifications among incidents classified by detailed incident form as completed rape, by sample type and mode for females, 2014–2015

BSQ ^c classification	DIF ^a classification: Completed rape ^b			
	General population		Volunteer sample	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Completed rape ^b	69.4 %	64.6 %	70.2 %	77.1 %
SV 1: Forced vaginal sex	27.8	18.8	26.5	27.5
SV 2 – SV 4: Forced oral, anal, or digital penetration	18.3	26.6	18.9	27.5
SV 5: Penetration while unable to consent	23.4	19.1	24.8	22.0
Other unwanted sex against will ^f	15.3 %	20.0 %	16.4 %	10.1 %
Other sexual contact/non-contact against will ^g	15.3 %	15.4 %	13.4 %	12.8 %
Number of unweighted sample cases	104	76	238	109

Note: General population estimates are based on weighted data for ages 18-49. Volunteer sample estimates are based on unweighted data for ages 18-29. See Appendix A for standard errors.

* Comparison group.

^aDetailed incident form questions used to classify the incident.

^bIncludes penetrative sexual contact using force or while unable to consent where the offender completed the behavior.

^cBehaviorally specific questions used in the sexual victimization screener questions.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fIncludes sexual victimization screening items about coerced penetration, other unwanted penetration, and attempted penetration.

^gIncludes sexual victimization screening items about unwanted kissing, groping, attempted kissing or groping, sexual exposure, and participation in sexual photos or movies.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Investigating the Discrepancies Between BSQ and DIF Classifications for Completed Rape

To get a better understanding of what may be related to the above discrepancies, the narratives provided at the end of each incident form were examined for those incidents where the BSQ and the DIF classification did not agree. The narratives were of variable quality and were not available for all incidents. The discussion below is based on those incidents for which it was possible to judge whether either the BSQ and/or the DIF conflicted with what was on the narrative.

BSQ and DIF both point to completed rape

There were a total of 259 incidents where the BSQ and DIF both pointed to a completed rape and a narrative was provided. Among these, there were 79 incidents in which the narrative did not provide enough details to be able to determine if the incident was a completed rape. Of the remaining 180 incidents, there were 18 incidents in which the narrative conflicted with the BSQ and DIF as indicating a completed rape. In these instances, either the behavior or tactic described in the narrative did not meet the criteria defining the incident as a rape. In 162 of these incidents, the BSQ, the DIF, and the narrative consistently pointed to a completed rape. All of these incidents, as described in the narratives, meet the definitions intended by the survey.

Example 1: "... He ...wanted to have sex. I didn't want to ... He pulled out his private parts and tried to force me to give him oral sex ...he pinned me down and did ... sex with me." (BSQ = forced vaginal sex; DIF = forced vaginal, oral, anal, digital penetration)

Example 2: "I was trying to sleep and the person woke me up and forced me to suck their penis. I refused, they forced me down and forced their penis in my mouth. He held me down, forced my head in certain movements." (BSQ = forced oral penetration; DIF = forced oral penetration;)

Example 3: "He just pinned me down and pushed me on the bed and ... he tried to kiss me. He tried to take my pants off and he did stick his fingers inside me. But I told him to stop and to get off. I kept telling him to stop and get off ... he did get off." (BSQ = forced digital penetration; DIF = forced digital penetration)

Notably, in some of these incidents, the BSQ and DIF both indicate a completed rape, but the BSQ points to a different type of sexual penetration than the DIF. Ultimately, the classification of completed rape is correct, but the three sources are not in identical alignment.

Example 4: "He took me to his ... house ... but because I could barely walk he told me I could lie down on the couch.... And he left then came back and tried to kiss on me. I tried to push him off, He ...forced himself inside me. I kept telling him to stop but he wouldn't." (BSQ = forced oral sex; DIF = forced vaginal penetration)

Example 5: "...He was touching my breasts and my butt and I told him to stop but didn't ... he went into my pants and put his fingers in my vagina. I was struggling severely at that point." (BSQ = forced vaginal penetration; DIF = forced digital penetration, threatened vaginal and oral penetration)

BSQ is completed rape and DIF is not a completed rape

There were a total of 114 incidents with a narrative where the BSQ indicated a completed rape and the DIF did not. Some of these narratives lacked the necessary detail about what happened, with 56 out of the 114 providing sufficient detail to be able to determine whether the BSQ or DIF was inaccurate. These are described below.

BSQ is Not Correct. For 34 of the 56 incidents, the narrative indicated that the BSQ was not reflecting key information described in the narrative. In more than half (21) of these narratives, the BSQ indicated a completed rape, but according to the narrative, the incident was more likely an attempted or threatened rape, rather than a completed one. A few examples from the narratives are provided below.

Example 6: "I opened the door ... but he slammed it back closed and locked it.... Then my friend came to the door ... and that's when he eventually let me out." (BSQ = completed oral sex; DIF = threatened forced vaginal penetration, attempted forced oral sex)

Example 7: "He tried to force himself on me. He said if I didn't do it he would find someone else to have sex with and leave me. He put his penis in my face. I got mad and he stopped." (BSQ = penetration while unable to consent; DIF = attempted forced penetration)

The reason for the differences may be related to the order of the BSQ questions. The first seven BSQ questions ask about completed acts, while the eighth item (SV8) is the first to ask about threats or attempts. Some respondents may be reporting on the attempts in response to the question about completed penetration, without knowing that a question about attempts is forthcoming.

Many of the remaining 13 incidents of completed rapes by the BSQ but not a rape on the DIF were unwanted contacts or attempts that were stopped by the victim. The respondents do not indicate any use of force in the narrative but do indicate an attempted sexual contact, as shown below.

Example 8: "He tried to start sex and I didn't want it. I said no, he stopped." (BSQ = forced vaginal penetration; DIF = unwanted threatened vaginal penetration and oral sex;)

Example 9: “A man tried to kiss and grope me and I shrugged him off.” (BSQ = penetration while unable to consent; DIF = attempted unwanted kissing, threatened groping;)

In both of these cases, the respondent did not report any force being used to carry out the incident on the DIF. This is why the incidents are classified as “unwanted” on the DIF rather than a sexual assault.

DIF is Not Correct. For 22 of the 56 narratives, the narrative was not consistent with the DIF classification. Among these, some cases reflect inconsistencies where the narrative and the BSQ both indicate a completed act, while the DIF points to an attempted or threatened act. Several examples of this follow below, which all involve digital penetration.

Example 10: “(He) pinned me against the ...wall, and penetrated me with his fingers.” (BSQ = forced digital penetration; DIF = attempted forced digital penetration).

Example 11: “I was at work and a man put his fingers in my vagina against my will.” (BSQ = forced digital penetration; DIF = attempted coercive digital penetration;).

Example 12: “We kept kissing but he still wouldn’t let me get up off the couch. I blacked out for a little bit and next thing I know we’re having sex.” (BSQ = sex when victim was unable to consent; DIF = attempted forced digital penetration;)

These discrepancies may be related to measurement error in DIF items asking the respondent to clarify the type of behavior that was used in the incident. In the DIF, respondents were offered four response categories for each behavior in order to classify the behavior as threatened only, attempted only, completed, or it did not happen. Some respondents may have gotten confused when making these distinctions.

Fourteen of the 22 instances where the DIF was inconsistent with the BSQ and narrative involved respondents who selected SV5 (sex when unable to consent) on the BSQ. Several of these were instances when the behavior was described as attempted/threatened on the DIF but completed in the narrative (see discussion above). Four others were instances where the DIF did not capture that the respondent was unable to consent. This was due to respondents skipping over many of the items related determining alcohol use because they reported that they could not remember any part of the incident.

Several others who selected SV5 did not report any penetrative behavior on the DIF but did describe this behavior in the narrative. One example is shown below.

Example 13: "It was a video-recorded situation. I was passed out drunk and my boyfriend had vaginal sex with me while I was passed out. No weapons were used." (BSQ = sex when victim was unable to consent; DIF = exposure when victim was unable to consent;).

In this case, the respondent reported on the DIF that she was made to participate in a sexual movie, but she did not indicate any penetrative behaviors on the DIF to correspond with the narrative or the BSQ.

A similar example is shown below.

Example 14: "We drank a little too much and got drunk. We ended up back at his place where we had oral and vaginal intercourse that I was in and out of being blacked out drunk the entire time. ... I wouldn't want to do it if I was sober, I was inhibited by alcohol. I made a poor decision." (BSQ = sex when victim was unable to consent; DIF = no unwanted behavior;).

The respondent selected the BSQ indicating sex when she was unable to consent, but did not indicate on the DIF that any of the unwanted behaviors occurred. Nor did she report in the DIF that she was unable to consent because of alcohol or drugs.

In the DIF, respondents were asked to focus on aspects of the incident that were “unwanted.” It may be that these respondents who screened in as having had sex when unable to consent, but indicated no unwanted behavior in the DIF, did not think of the behavior as “unwanted”⁴⁴. In the above examples, respondents describe situations of voluntary use of alcohol and the respondents seem to be hesitant to describe the behaviors as “unwanted.” This is best illustrated in Example 13 above, where the only behavior that was described this way on the DIF was the videotaping of the event.

DIF indicates a completed rape and BSQ does not

There were 110 incidents with narratives where the DIF classified the incident as a completed rape but the BSQ did not classify it this way. In 54 of these incidents, respondents provided narratives with enough information to compare to the BSQ and DIF classifications. In 51

⁴⁴It is also possible that these respondents did not remember. But they did report this in SV5 and said “no” when asked directly to D1 on the DIF.

of these 54 incidents, the narratives were inconsistent with the BSQ item, and in 3 of the cases, the narrative was inconsistent with the DIF.

BSQ is not correct. Among the 51 incidents in which the narrative was inconsistent with the BSQ item, 26 were incidents in which the respondent indicated one of the BSQ items SV9 – SV14 but the narrative indicated a completed rape. Items SV9 – SV14 are intended to measure non-penetrative sexual contact and non-contact. Several examples of this type of discrepancy follow.

Example 15: "I felt his hand go inside my pants.... The next thing I know is he tried to force himself on me.... After the incident happened, he through (sic) my clothes at my face and told me to clean myself up." (BSQ = fondling, groping, grabbing, touching; DIF = forced digital penetration;)

Example 16: "My ex ... threatened me.... I resisted, she kissed me, held my face. And pinned me to the bed, performed oral sex and left." (BSQ = fondling, groping, grabbing, touching; DIF = forced digital penetration;)

Example 17: "He grabbed me and he began to kiss me. I told him, 'No...' He pinned me to the bed andput his hands inside of me." (BSQ = Attempted kissing, fondling, groping, grabbing, touching; DIF = forced digital penetration;).

Example 18: "He pushed me down on the couch and proceeded to tear my clothes off and ...he pushed my legs open and inserted." (BSQ = exposure; DIF = forced oral sex, anal sex and digital penetration;)

Of the remaining incidents where the BSQ was not consistent with the narrative, 12 were for BSQs that asked about penetration involving tactics other than force or inability to consent. SV6 refers to coercion and SV7 refers to any other situations where the respondent was made to have sex against her will. These narratives generally were consistent on the behavior (penetration), but were not consistent with the tactics indicated in the narrative. An example is shown below.

Example 19: "I could (sic) not move. ... When he was done he left I could not move friends found me outback (sic)." (BSQ = coerced penetration; DIF = forced vaginal and digital penetration;).

Finally, there were 13 incidents in which the BSQ described attempted, but not completed, penetration against the respondent's will (SV8), but the narrative indicated that penetration was completed. The next section provides a fuller accounting of the incidents reported for SV8.

DIF is not correct. There were three incidents in which the narrative suggests that the DIF may be incorrect. In these cases, the BSQ referred to sexual assaults involving kissing, groping, and

other sexual activities (SV9 – SV14) but the DIF included some type of completed penetration. Two of these examples are provided below.

Example 20: "...it was ... rush hour...he just grabbed at me and tried to kiss me and then ... he just jump off the train like it was a joke." (BSQ = fondling, groping, grabbing, touching; DIF = forced vaginal penetration;).

Example 21: "I was at work and a coworker came up and gave me a hug ... and then he kissed me on the neck and then a little bit later in the day he put my hand on his crotch, and then he blocked my way so that I couldn't go where I needed to go, and I said I have work to do and he moved." (BSQ = kissing; DIF = forced vaginal penetration;).

In both of these cases, the respondent answered that completed penetration occurred on the DIF.

Both of the above interviews were conducted over the telephone and it was possible to listen to the audio recordings. For Example 20, the interviewer did not clearly understand what the respondent was saying. The respondent initially answered “no” to the question of whether there was penetration, but once the interviewer read the response categories, the respondent selected the wrong one (using a number to signify the answer). For Example 21, the interviewer stumbled across various parts of the question on penetration. Once the respondent answered the question, she indicated that she was saying “yes” that she was groped, but did not mention penetration. The respondent seemed to have lost track of the initial part of the question, which asked about penetration, and concentrated on what she had initially reported happened.

This points to difficulties administering the DIF questions on unwanted behavior over the telephone. Other evidence of this can be seen in section 11.3.1, which discusses the interviewer behavior on the CATI for these items. The interviewers were more likely to stumble on these questions and the initial response categories. Furthermore, respondents tended to answer the initial portion of the question with a “yes/no” (“Did the person threaten, try to or actually put his penis in your vagina when you didn’t want it to happen?”), rather than waiting for the interviewer to read the four different response categories. This interrupted the flow of the question and led to some confusion.

Discussion of Completed Rape Measurement

Overall, the narratives indicate that most incidents that were classified by the BSQ or DIF as a completed rape can reasonably be considered a completed rape. Very few of the incidents that had a complete narrative described instances that did not involve force or while unable to consent. There were a few instances where the incident did not meet the legal criteria—for example, no force was applied or the behavior was not considered unwanted (e.g., when alcohol was involved), but these were more the exception, rather than the rule.

While the prevalence rates based on the BSQs and DIF are very similar, there were differences with respect to which incidents were classified into the completed rape category. Between 25 percent and 50 percent of the incidents screened in as a completed rape using the BSQs were not classified as such using the DIF. About 30 percent of the incidents that screened in as something other than a completed rape were classified by the DIF as a completed rape.

The above analysis relied on the narratives to reconcile the discrepancies between the BSQ and the DIF. The narratives are of varying quality with respect to enumerating the particular behaviors, tactics, and levels of consent. For this reason, it is difficult to estimate the percentage of false positive and false negatives related to either the BSQs or the DIF. However, the above analysis provides evidence that there is measurement error for both types of questions and points to the sources of this error.

With respect to the BSQs, the errors found were consistent with previous discussions (e.g., Steiger et al., 2014; Cantor & Steiger, 2014; Cook et al., 2011). The BSQs pack a lot of information into a single question, covering the requisite behaviors, tactics, and consent condition. Some respondents may answer affirmatively to one of these items if their experience includes some, but not all, elements of these conditions (false positives). For example, many of the discrepancies were found to be attempts or threats as described in the narratives and the DIF, rather than completed rapes as indicated by the BSQ. Others included unwanted contacts that did not seem to contain the tactic (e.g., force), or it was stopped by the victim before any force was applied.

There were a number of respondents who gave an affirmative responses to SV9 – SV14 (sexual touching or exposure) and were found to meet the criteria for a completed rape. There are several reasons this may be occurring. One is that respondents may not initially remember a particular event when the first BSQ is administered. Once administering other BSQs covering related behaviors, the respondent may remember the event after thinking about it more and/or

hearing about something that reminded her of the incident. A second possible reason is that respondent may be undecided whether she wants to report a particular event at the point of the screening question. Once going through more of the screener, she may change her mind and report it at a later question.

For a two-stage design, the fact that these incidents were not captured at the designated screening question may not be problematic as long as all incidents are captured by the time the full screener is administered. However, it is important to clarify the screening questions when respondents are confused. For example, there was some evidence that relying on respondents to recall the definition of “sex” when answering later questions (e.g., SV6, SV7, SV8) may not have worked for all respondents. This seemed most apparent in the prevalence of reports of completed digital penetration on the DIF being reported in response to the attempted rape item (SV8). Some respondents may not have thought of digital penetration as the type of sex referenced in this BSQ. Future implementation of the BSQs should consider wording that keeps the definition available to the respondent.

One source of error related to the DIF was centered on the unwanted behavior questions. These questions attempted to make distinctions between threats, attempts, and completed acts. This level of detail may have confused some respondents when trying to describe what happened. This was especially evident for the phone interviews, where, in several cases, the respondent became confused and the interviewer recorded the wrong information. Future implementation of a DIF should simplify these questions, perhaps dropping the distinction between threats and attempts (see also discussion of attempted rapes in section 9.3.2).

A second issue that revolved around the unwanted behavior items was the tendency for ACASI respondents to not report any unwanted behavior. It is unclear what is behind this mode effect. Some of these were found to come from the BSQ that related to sex while unable to consent (SV5). Respondents were more likely to say on the ACASI that the behavior did not qualify as “unwanted.” Future implementation of a two-stage design might consider checking back with any respondent who does not report any unwanted behavior to make sure she had understood the questions as intended. A third issue revealed by the above comparison are difficulties capturing whether the victim was unable to consent when respondents cannot remember what happened. There were a small number of incidents where the narrative indicated the respondent was either blacked out or unconscious during some or all the incident. However, when answering the questions on the DIF, they did not get all of the questions that determine inability to consent because she could not remember any part of the incident (there were four individuals who fit this profile). Some

consideration should be given to including those who cannot remember what happened as unable to consent.

9.3.2 Comparison of Classifications for Attempted and Threatened Penetration

In this section the correspondence between the BSQs, DIF, and narratives are discussed for attempted and threatened acts.

One of the BSQ screening items (SV8) asks if anyone “tried, but did not succeed” at making the woman “have any type of sex against her will.” These responses are compared to the DIF responses that involved force, inability to consent, coercion or other unwanted behavior. Note that this is not entirely comparable, as the BSQ refers to “tried but did not succeed” but does not define the specific tactic (e.g., physical, inability to consent, non-physical threat). The DIF items specifically make this distinction.

Table 9-11 provides the prevalence rates from the BSQ versus the classification based on the DIF. Using the BSQs, the past 12-month rate of attempted penetration is 3.5 percent in ACASI and 2.2 percent in CATI. Alternately, the past 12-month prevalence of attempted or threatened penetration using any tactic in the DIF is 3.0 percent in ACASI and 2.3 percent in CATI. Note that the BSQ only mentions “attempted” penetration, whereas the DIF takes into account “threatened” penetration. If we had included only attempted penetration for the DIF, the rates would have declined slightly.

Table 9-11. Prevalence rate of attempted penetration by mode of interview and classification method for females ages 18-49 in the general population, 2014–2015

Classification	ACASI ^a		CATI ^b	
	BSQ ^c	DIF ^{d*}	BSQ ^c	DIF ^{d*}
Attempted unwanted penetration ^e	3.5 %	3.0 %	2.2 %	2.3 %

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

^cBehaviorally specific questions used in the sexual victimization screener questions.

^dDetailed incident form questions used to classify the incident.

^eIncludes penetrative sexual contact by any tactic where the offender physically tried, but did not complete the behavior.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Attempted and Threatened Penetration on the BSQ

While the differences in prevalence rates between the BSQ and DIF are not significantly different, there is significant movement between categories across the two methodologies (table 9-12). For the general population sample, roughly half of incidents that screened in as penetration that were tried but not successful were not classified by the DIF as attempted or threatened penetration. ACASI incidents were significantly less likely to have been classified consistently by the BSQ and DIF (44.5%) than CATI incidents (51.3%).

For incidents reported at SV8 on the ACASI, more than one-third were ultimately classified as completed or attempted sexual touching on the DIF. In CATI incidents, this rate was slightly lower, although still more than 27 percent were classified as touching. As observed with the comparison of completed rapes between the BSQ and the DIF, ACASI incidents were significantly more likely to not have been classified into a category because the respondent did not report any unwanted behavior or tactic information to classify the event. The rates are 9.3 percent for ACASI incidents and 1.0 percent for CATI incidents.

The results for the VO sample are similar to that of the general population, although the differences between ACASI and CATI are not significant.

Table 9-12. Distribution of detailed incident form classifications among incidents classified by behaviorally specific questions as attempted penetration, by sample type, and mode for females, 2014–2015

DIF ^c classification	BSQ ^a classification: Attempted penetration against will ^b (SV8)			
	General population		Volunteer sample	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e
Attempted and threatened penetration ^f	44.5 %†	51.3 %	45.5 %	56.0 %
Attempted rape ^g	19.5	20.6	17.5	26.8
Other attempted unwanted penetration ^h	6.1 †	20.4	9.7	8.5
Threatened rape ⁱ	7.9	6.2	8.6	11.0
Other threatened unwanted penetration ^j	11.0	4.1	9.7	9.7
Completed rape ^k	5.7 %	13.6 %	10.4 %	6.1 %
Other completed unwanted penetration ^l	2.7 %	6.8 %	2.6 %	7.3 %
Completed unwanted sexual touching ^m	21.5 %	20.4 %	19.7 %	13.4 %
Attempted and threatened unwanted sexual touching ⁿ	16.1 %	7.1 %	10.0 %	7.3 %

Table 9-12. Distribution of detailed incident form classifications among incidents classified by behaviorally specific questions as attempted penetration, by sample type, and mode for females, 2014–2015 (continued)

DIF ^c classification	BSQ ^a classification: Attempted penetration against will ^b (SV8)			
	General population		Volunteer sample	
	ACASI ^d	CATI ^e *	ACASI ^d	CATI ^e
Not enough information ^o	9.3 %†	1.0 %! ^f	11.6 %	9.8 %
No unwanted behavior	6.1 %	1.0 %!	8.6 %	8.5 %
Not enough behavior information to classify ^p	2.1 %!	--	2.2 %	--
Not enough tactic information to classify ^q	1.1 %!	--	0.7 %!	1.2 %!
Number of unweighted sample cases	139	85	279	92

Note: General population estimates are based on weighted data for ages 18-49. Volunteer sample estimates are based on unweighted data for ages 18-29. See Appendix A for standard errors

*Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

--Less than 0.05%.

^aBehaviorally specific questions used in the sexual victimization screener questions.

^bIncludes unwanted penetrative sexual contact, not including force, while unable to consent, or coercion.

^cDetailed incident form questions used to classify the incident.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fIncludes penetrative sexual contact using any tactic where the offender verbally threatened or physically tried, but did not complete the behavior.

^gIncludes penetrative sexual contact using force or while unable to consent where the offender physically tried, but did not complete the behavior.

^hIncludes penetrative sexual contact where the offender physically tried, but did not complete the behavior and force or while unable to consent was not reported.

ⁱIncludes penetrative sexual contact using force or while unable to consent where the offender verbally threatened, but did not physically attempt the behavior.

^jIncludes penetrative sexual contact where the offender verbally threatened, but did not physically attempt the behavior and force or while unable to consent was not reported.

^kIncludes penetrative sexual contact using force or while unable to consent where the offender completed the behavior.

^lIncludes penetrative sexual contact where the offender completed the behavior and force or while unable to consent was not reported.

^mIncludes unwanted non-penetrative sexual contact using any tactic where the offender completed the behavior.

ⁿIncludes unwanted non-penetrative sexual contact using any tactic where the offender verbally threatened or physically tried, but did not complete the behavior.

^oIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all behavior and tactic items.

^pIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all behavior items.

^qIncludes incidents where respondent said “no,” refused to answer, or didn’t know the answer to all tactic items.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Attempted or Threatened Penetration on the DIF

When examining incidents that were classified as attempted or threatened penetration by the DIF, slightly less than half of them screened in as something other than attempted penetration (table 9-13). In the general population sample, 55 to 59 percent of incidents ultimately classified as attempted or threatened penetration also screened in as attempted penetration on the BSQ. Additionally, 15 to 22 percent screened in as a completed rape (SV1 through SV5), and 18 to 22 percent screened in as sexual contact. There were no differences by mode of interview.

For volunteer (VO) respondents, ACASI attempted or threatened penetration screened in as attempted penetration at a similar rate to the general population sample (61.0%). CATI volunteers were significantly less consistent between the BSQ and the DIF classification, with 42.6 percent of attempted or threatened penetration screening in as attempted penetration.

Table 9-13. Distribution of behaviorally specific question classifications among incidents classified by detailed incident form as attempted penetration, by sample type, and mode for females, 2014–2015

BSQ ^c classification	DIF ^a classification: Attempted penetration against will ^b			
	General population		Volunteer sample	
	ACASI ^d	CATI ^{e*}	ACASI ^d	CATI ^{e*}
Completed rape ^f	21.8 %	15.0 %	16.5 %	17.6 %
Other completed unwanted penetration ^g	3.9 !	3.7 !	5.5	8.3
Attempted unwanted penetration ^h	55.6	59.1	61.0 †	42.6
Sexual contact ⁱ	18.7	22.2	17.0 †	31.5
Number of unweighted sample cases	103	76	200	108

Note: General population estimates are based on weighted data for ages 18–49. Volunteer sample estimates are based on unweighted data for ages 18–29. See Appendix A for standard errors

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

^aDetailed incident form questions used to classify the incident.

^bIncludes penetrative sexual contact using any tactic where the offender physically tried, but did not complete the behavior.

^cBehaviorally specific questions used in the sexual victimization screener questions.

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fIncludes sexual victimization screening items about vaginal, oral, anal, and digital forced penetration and penetration while unable to consent due to alcohol or drugs.

^gIncludes screening items measuring penetration using coercion and other unwanted penetration.

^hIncludes screening item measuring penetrative sexual contact using any tactic where the offender physically tried, but did not complete the behavior.

ⁱIncludes sexual victimization screening items about unwanted kissing, groping, attempted kissing or groping, sexual exposure, and participation in sexual photos or movies.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

Investigating the Discrepancies Between BSQ and DIF Classifications for Attempted Penetration

The narratives provided at the end of each incident form were examined for those incidents where the BSQ and the DIF classification did not agree on the classification as an attempted or threatened penetrative act. The discussion below is based on those incidents for which it was possible to judge whether either the BSQ and/or the DIF conflicted with what was on the narrative.

BSQ and DIF both point to attempted or threatened penetration

There were a total of 125 incidents in which the BSQ and the DIF both pointed to attempted or threatened penetration. However, the narrative is not always in agreement with the BSQ and DIF classifications. Setting aside the 29 incidents in which the narrative does not contain enough information, 96 narratives are reviewed in the discussion below.

In 40 of the 96 incidents, the narrative was inconsistent with either the BSQ or the DIF. Many of these are situations in which the woman may have perceived the person's actions to be an attempt at sexual penetration, but the narrative does not contain any information pertaining to the initiation or threat of forcing sexual penetration. This may be due to an incomplete narrative, although these narratives seem to contain relatively complete information. It is more likely because the woman thought the kissing, groping, or sexual talk was an attempt to have sex.

Example 22: "I was at a club ... one of the guys ... grabbed me and pushed me against a wall, trying to kiss me. When I moved and told him to stop he said sorry and let me go." (BSQ = Attempted penetration; DIF = Attempted forced vaginal penetration)

Example 23: "Then he started kissing me, and I kissed back for a few seconds because I was too intoxicated to fully understand what was going on. I quickly pushed him away and told him no, then he tried kissing me again and I pushed him off." (BSQ = Attempted penetration; DIF = Threatened vaginal penetration while unable to consent)

Example 24: "She started groping me and I told her to stop ... she tried to kiss me and then ... I was able to get out of that situation." (BSQ = Attempted penetration; DIF = Threatened vaginal penetration, attempted digital penetration using force)

In each of these incidents there was clearly unwanted sexual touching (e.g., kissing, groping), but there was not an explicit physical attempt of penetration or explicit threat to do so.

In other incidents where the narrative is inconsistent with the BSQ and/or DIF, penetration actually occurred, but the woman may have only been thinking about whether vaginal penetration was attempted, even though some other type of penetration was completed.

Example 25: "He stuck his hand in my pants, he grabbed my breast, he pulled my pants down licked my vagina ... tried to have sex with me." (BSQ = Attempted penetration; DIF = Attempted vaginal, oral, digital penetration)

Example 26: "He began kiss me forcefully; I told him to stop, but then he start to pull down my pants ... Then he pin me down and began to give me oral ..." (BSQ = Attempted penetration; DIF = Attempted vaginal, oral, anal penetration)

For the remaining 56 incidents, the BSQ, the DIF, and the narrative are all in alignment, providing good examples of attempted penetration.

Example 27: "I gave a guy ... a ride home ... and he ... shoved his hands between my leg" (BSQ = Attempted penetration; DIF = Attempted vaginal and digital penetration with force)

Example 28: "A drunk guy ... tried to rape me and held my arms but I got loose and punched him in his face ... and I ran away." (BSQ = Attempted penetration; DIF = Attempted vaginal penetration with force)

Example 29: "I ... was locked in his room and he was trying to force me. I said 'no,' and, ...he got a little violent, and my friend came in and stopped him." (BSQ = Attempted penetration; DIF = Attempted vaginal and digital penetration with force)

BSQ points to attempted penetration, DIF does not

There were 340 incidents in which the BSQ indicated attempted penetration but the DIF pointed to some other classification. For 226 of these incidents, the respondent provided a complete enough narrative to be able to evaluate the discrepancy between the BSQ and the DIF.

Attempted vs. Completed Penetration. For 58 of these 226 incidents, the respondent described completed penetration in the DIF, rather than attempted or threatened penetration. In 17 of these 58 cases, the narratives suggest the BSQ classification of attempted penetration is more accurate than the DIF classification of completed penetration. Some examples are shown below.

Example 30: "... he tried to put his fingers in my vagina and I was kind of not wanting that and he kept trying to go for that and then he was groping me. ... I was able to get away from that situation." (BSQ = Attempted penetration; DIF = completed digital penetration)

Example 31: "...He unzipped his pants and thought we were going to have sex.... I kept saying 'no, that's not going to happen' ... he just walked away and that was that. We went home." (BSQ = Attempted penetration; DIF = completed oral penetration)

In 25 of these 58 cases, the narratives clearly suggest the DIF classification of completed penetration is more accurate than the BSQ classification of attempted penetration. Some examples are shown below.

Example 32: "He took off all his clothes ... then he pushed me down onto the bed and he didn't stop when I said stop ...and he eventually just passed out during the act. It lasted I wanna say about 45 minutes." (BSQ = Attempted penetration; DIF = completed anal, digital penetration)

Example 33: "We were in my bed and he was fingering me and I told him to stop and he wouldn't and then I screamed to stop and then he finally did...." (BSQ = Attempted penetration; DIF = completed digital penetration)

In the remaining 16 of these 58 cases, the narratives do not provide enough detail to determine whether the BSQ or the DIF classification is more accurate.

Attempted Penetration vs. Sexual Contact. For 137 of the 226 incidents, the respondent described completed, attempted, or threatened sexual contact on the DIF, rather than attempted or threatened penetration as indicated on the BSQ. In these situations, the woman may have been interpreting the person's advances as attempted sex when answering the BSQ, but in many cases, did not indicate any attempted penetrative behaviors in the DIF.

Example 34: "I was out at a bar ... and the guy next to me ... kept pushing me. ...I like kind of end up like running away." (BSQ = Attempted penetration; DIF = Attempted kissing, groping)

Example 35: "We were ... kissing. But then he kept trying to take off my clothes.... And I kept saying to him, 'No.' ... I kept having to say, 'No' until I finally left." (BSQ = Attempted penetration; DIF = Completed groping)

Example 36: "We were drinking ... she kissed me and grabbed my butt. I told her to stop. She stopped...." (BSQ = Attempted penetration; DIF = Completed kissing, groping)

In other examples, the woman did describe behaviors that were more aligned with the BSQ classification of attempted sex, rather than the DIF classification of sexual contact.

Example 37: ‘He pinned me ...and took off my shirt and tried to take off my bra. He ...pinned me to the bed ... he was saying something like ... he wouldn’t let me move. Then he kept trying to kiss me and I kept telling him that I didn’t want to do anything.... I must have just resisted him and then left. (BSQ = Attempted penetration; DIF = Completed kissing, groping)

Example 38: He began to try to kiss me and things. He started to unbutton my jeans. I was screaming no and stop. He was on top of me and I couldn’t get him off ... Then he got off. (BSQ = Attempted penetration; DIF = Completed kissing, groping)

Attempted Penetration vs. no Unwanted Behavior. For 31 of the 226 incidents, the respondent did not report any unwanted behavior in the DIF, even though she reported attempted or threatened penetration on the BSQ. In most of these cases, the narrative is consistent with the DIF in that no attempted penetration is described. Some examples are shown below.

Example 39: ‘I was sitting at the bus stop and a man in a car pulled up and offered me a ride. ... Five minutes later the man came back ... minutes later he came back (again).’ (BSQ = Attempted penetration; DIF = No behavior reported; no force, coercion or inability to consent)

Example 40: ‘Someone called me threatening to post nude photos of me online.... I ignored his request ...’ (BSQ = Attempted penetration; DIF = Unknown behavior; no force, coercion or inability to consent)

DIF Points to Attempted Penetration, BSQ Does Not

There were 234 incidents, of which 159 provided a complete narrative in which the DIF classification was attempted or threatened penetration, but the BSQ response was something other than attempted penetration.

Attempted vs. Completed Penetration. For 77 of these 159 incidents, the BSQ indicated completed penetration, rather than attempted penetration. Among these 77 incidents, roughly half had narratives suggesting that the BSQ was accurate, but the DIF classification was not. Some examples are shown here.

Example 41: ‘He started lifting up my clothes.... Eventually he got my clothes off and he penetrated, and then that was it. I don’t remember what happened after that.’ (BSQ = Completed penetration while unable to consent; DIF = Attempted oral penetration)

Example 42: "...He pinned me against the elevator wall, and penetrated me with his fingers...." (BSQ = Completed forced digital penetration; DIF = Attempted vaginal and digital penetration)

For the other half of the 77 incidents, the narrative suggested that the DIF classification was more accurate than the BSQ, indicating attempted, rather than completed, penetration.

Example 43: "... he grab my head and press it down in his penis ... one of his friends came, and he heard it. That's when he stopped ... and left." (BSQ = Completed oral penetration; DIF = Attempted oral penetration while unable to consent)

Example 44: "I was walking down the street ... and I seen a guy that looked like he was a little bit under the influence. ... And so he ran up to me and ... tried to stick his tongue in my mouth." (BSQ = Completed penetration while unable to consent; DIF = Threatened vaginal, oral, digital penetration)

Attempted Penetration vs. Sexual Contact. In 82 of the 159 incidents, the DIF indicated attempted or threatened penetration, but the BSQ screening item indicated some type of unwanted sexual contact, such as kissing, groping, or exposure. Of these 82 incidents, 57 are instances where there could have been either an attempt or threat of some type of penetration, but the description is not specific enough. However, if one assumes these narratives are complete, they indicate that when answering the DIF as an attempt or threat, the respondent was thinking of what may have happened if the situation was allowed to go further than it did.

Example 45: "I fell asleep on his couch, ... and he started to try groping and touching I said stop again and I got up and yelled at him and left." (BSQ = Attempted kissing or groping; DIF = Attempted vaginal, digital penetration)

Example 46: "A large man ... physically attempted to block me from leaving. ... I ran away outside." (BSQ = Completed groping; DIF = Attempted vaginal penetration)

In the other 25 of these 82 cases, the DIF classification of attempted or threatened penetration seems to be more of an accurate characterization of the incident based on the narrative than the BSQ sexual contact.

Example 47: "He tried to push me down on the couch.... I don't really recall what exactly happened. ... I ended up pushing him off the couch. ... I ran out of the house and went home." (BSQ = Attempted groping; DIF = Attempted digital penetration)

Example 48: "A person ... exposed himself ... He then proceeded to push me to the floor to try to have intercourse. He kissed me, pushed him off... Heproceeded to try to penetrate and pull my clothes off and all that kind of stuff. That's when I pushed him off." (BSQ = Exposure; DIF = Attempted vaginal, oral, digital penetration)

Example 49: “He came up behind me and started groping me--attempted to penetrate my vagina... I ... pushed him away, he attempted again and I ... was able to push him away again. ...He sort of went off without further incident...” (BSQ = Attempted groping; DIF = Attempted digital penetration)

Discussion of Attempted Penetration Measurement

The analysis presented in this chapter reveals the complexity of measuring attempted or threatened penetration. From a legal point of view, a threat is an action that indicates an intent without actually trying to do it (e.g., verbally threatening). The action needs to be backed up by some credible threat that it can be carried out (trapping someone in a car or room; physical intimidation). An attempt requires that the offender take a “substantial” step toward completing the behavior. Respondents’ perceptions of a threat or an attempt did not necessarily line up with these legal criteria. There were a number of discrepancies where the narrative indicated that the incident involved non-penetrative sexual contact, which the respondent seemed to interpret as a threat or attempt to go further. There was no indication of either a verbal statement of intention or physical actions to go further. There were also situations where it wasn’t clear there was a credible threat – particularly groping incidents that occurred in a public place.

By following up with specific questions about the behaviors, the DIF is intended to provide the detail needed to classify the event. As noted above, there were a number of examples where this worked as intended and the respondent reported an attempt or threat in the BSQ, but later said it was sexual contact, attempted sexual contact, or no unwanted behavior (e.g., examples 34, 36). But there were also cases that seemed to be aggressive sexual contacts with no outward attempt at penetration but were reported as attempts or threats at penetration on the DIF (e.g., examples 22, 23, 46). Some of the problem is with the wording of the question on the DIF that asks about behaviors. For example, the question asking about unwanted vaginal penetration was

Did the person threaten to, try to, or actually put his penis in your vagina when you didn’t want it to happen?

- Yes, the person verbally threatened to do this but did not physically try to do it.
- Yes, the person physically tried to do this but did not actually do it.
- Yes, the person actually did it.
- No, this did not happen.

This format was used for each type of penetration (oral, digital, anal) and non-penetrative sexual contact (e.g., kissing, groping, exposure). This format used the response categories to define more specifically what was meant by attempt or threat (i.e., verbally threatened; physically tried to do this). However, this distinction may have been lost on some respondents. As a result, some respondents used a wider definition of “attempt” or “threat” than was intended. To improve the measurement of attempted or threatened behaviors, it is necessary to more clearly define what qualifies under these concepts. However, it should also be recognized that it is difficult for a survey to capture the entire context of the incident, including the intent of the perpetrator. In some of the examples cited above, the victim’s understanding of this context was that it was an attempt or threat at penetration. While the measurement used on the RSA Pilot Test can be improved, this has to be balanced against the extent that it is possible to make these distinctions on a survey. We discuss this in more detail in the final chapter on recommendations.

A second issue with the DIF items is that the tactic questions (type of force used or threatened) are not specific to particular types of behaviors. Questions D1 and D2 ask about the unwanted behaviors and questions D3 and D4 ask about the particular tactics. However, in situations where there is unwanted sexual contact (from D2) and an attempt/threat of penetrative contact (From D1), it is not possible to assess whether the threat or attempt was associated with the force or threat of force. We discuss possible solutions to this issue in the final chapter as well.

10. Nonresponse Bias and Related Concerns

Nonresponse bias distorts the final estimates from a survey when systematic differences between the respondents and nonrespondents remain after weighting adjustments, so that the respondents do not adequately represent the target population as a whole. Because nonresponse is large in both the ACASI and CATI samples, there is a risk that nonresponse bias could substantially affect the findings. This chapter summarizes evidence on this question.

Because victimization rates for nonrespondents are unavailable, there is no direct answer to the question of how closely the RSA Pilot Test respondents resemble the nonrespondents on victimization. However, evidence from the weighting provides evidence on a related question: How similar are respondents and nonrespondents on the characteristics used in the weighting? As noted in Chapter 3, these characteristics are age, education, race/ethnicity, and marital status. If large differences appear for the weighting variables, there is some basis for concern, although not conclusive evidence, that nonresponse may affect the coverage of other characteristics, such as victimization. Conversely, if the comparisons imply that two samples seemed to have covered the target population relatively uniformly, the assessment is more favorable, but again not conclusive. The first section analyzes evidence on nonresponse and compares the final estimates from the two samples to estimates using intermediate weights during the weighting process.

A second analysis compares the neighborhood characteristics of respondents with those of nonrespondents. The comparison is relatively straightforward for the ACASI sample, because the sample addresses are known for both the respondents and nonrespondents, and neighborhood characteristics can be studied at the detailed level of the census tract. The comparison is problematic for the CATI sample, however, because the location of nonrespondents can be reliably inferred only at the county level, at best.

A third analysis concerns the degree to which the ACASI target population may have differed from the CATI target population. The ACASI target population was geographically defined by residence of the respondents, and the ACASI address-based sample accurately conformed to the boundaries of the CBSAs. With RDD samples of landline and cell phones, the CATI sample relied initially on ranges of telephone numbers—area codes and exchanges within them—to indicate whether the respondent resided in the CBSA. As a consequence, the CATI sample is less precisely defined geographically than the ACASI sample. For example, because of the portability of cell phone numbers, some CBSA residents have cell phones with phone numbers outside the set of area

codes from which the RDD samples were drawn. With respect to the CATI sample, these persons may be thought of as in-movers. Conversely, out-movers may have left the CBSAs but retained their cell numbers. The third section reports what can be learned about characteristics of in-movers and out-movers from the available RSA Pilot Test data.

A fourth analysis conducts a level-of-effort analysis for both the ACASI and CATI interviews. This analysis examines whether the rates of rape and sexual assault differ for those that were relatively easy to interview versus those that were harder to interview (or took more effort). This is one of several methods that can be used to assess nonresponse bias. It relies on the assumption that those who were the hardest to interview resemble those who were not interviewed at all.

10.1 Effects of Differential Response Rates

Chapter 3 describes the weighting of the ACASI and CATI samples. It includes comparisons of estimates of the control variables based on intermediate weights before raking with the corresponding distributions from the American Community Survey (ACS).

For the ACASI sample, the person-level base weights and final population weights estimate similar distributions for the control variables used in the raking adjustments: race/ethnicity, age, education, and marital status (tables 3-7 and 3-8). Hispanic and non-Hispanic black women are well represented overall, and the variation in the coverage from one CBSA to another may be due to segment-level clustering. Married women appear slightly overrepresented. The most notable deviation is for ages 18-29, where the person-level base weights underestimate the proportion in the population when compared to the ACS, both overall and in each of the CBSAs separately. The differences are most pronounced for ages 18-21, where the shortfall is almost 15 percent for the population weighted estimate (11.47% in the ACS vs. 9.81%). There is a similar shortfall of about 12 percent for ages 22-24. In other words, women in the highest risk age groups are underrepresented in the sample before the ACASI weights are adjusted to ACS totals.

For the CATI sample, the composite person weights and the final population weights estimate similar distributions for the control variables (tables 3-9 and 3-10). Again, Hispanic and non-Hispanic black women are well represented. Married women appear underrepresented, but college educated women are overrepresented. In contrast to the age distribution of the ACASI

sample, the youngest age group, 18-24, is overrepresented in the CATI sample by about 10 percent. In short, the CATI sample may have slightly overrepresented the women at highest risk.

The final weights for both the ACASI and CATI samples were constructed in a series of steps, producing intermediate sets of weights in the process. As a review of the weighting of the ACASI sample, the construction of the equalized final weights started from household-level base weights, given in Chapter 3 by equation (3.2):

- Each household-level base weight was the reciprocal of the probability of selection of the housing unit or college dormitory room. Except for eight segments noted in Chapter 3, these weights are available for the entire sample. In the absence of nonresponse, the household-level weights would theoretically produce unbiased estimates of household characteristics for the CBSAs.
- For households completing the screening questionnaire, person-level base weights (3.3) were computed by adjusting the household-level base weights for subsampling a single respondent among eligible respondents. In the absence of nonresponse, the person-level base weights would produce theoretically unbiased estimates of person-level characteristics for females ages 18-49 in households or college dormitories. The person-level base weights are available only for sample households with completed screening questionnaires.
- Within each CBSA, the person-level base weights of respondents completing the ACASI interview were then raked to totals derived from the ACS, to produce the population adjusted final weights (3.4).
- As a final step, these weights were then equalized by multiplying them by a single factor in each of the CBSAs in order to treat each CBSA as if it were the same size as the others, shown as (3.5).

Weighting for the CATI sample followed a similar broad outline but included additional steps:

- A screener-level base weight was created from the inverse of the sampling probability for the telephone number (3.6).
- The screener-level base weight was adjusted to remove telephone numbers found to be non-residential and to compensate for units with unknown residency status (3.7).
- A screener nonresponse adjustment was computed within cells determined by CBSA and phone sample type (3.8).
- Person-level base weights were created to reflect within-household subsampling when the screener identified more than one eligible respondent (3.9).

- An adjustment for nonresponse to the extended ACASI interview was then performed in a manner similar to the adjustment for screener nonresponse (3.10).
- The adjusted person-level weights for the landline and cell samples were then composited by a weighting factor of 0.5 for respondents represented in both frames; that is, when they had a cell phone for their own use but could also be reached by landline in their household.
- After combining the samples, the weights were trimmed to reduce extreme weights.
- Finally, the composite person-level weights were raked to control totals from the ACS (3.11).
- As a final step, equalized weights were produced in the same manner as for the ACASI sample (3.5).

Table 10-1 compares ACASI prevalence rates for rape and sexual assault with two sets of intermediate weights: the initial household base weight as the inverse of the probability of selection, and the person base weight incorporating an adjustment for subsampling of eligible respondents within a household. For each set of intermediate weights in the table, separate CBSA rates were computed and then averaged to simulate the effect of equalizing the final weights. Differences in the table using the household and person base weights are slight, but the raking step increases the estimates of key outcomes, some by about 10 percent. The increase is readily explained as the consequence of larger relative increases in the weights of the youngest age groups, who would be at the greatest risk, but the raking may have had other effects as well.

Table 10-1. Primary prevalence rates from ACASI samples using intermediate and final weights, 2014–2015

Characteristic	Household base weights ^a	Person base weights ^b	Final equalized weights ^c
Rape and sexual assault ^d	5.4 %	5.4 %	5.9 %
Completed ^e	4.1	4.1	4.4
Attempted ^f	1.5	1.6	1.8
Threatened ^g	0.6	0.6	0.7
Rape ^h	3.3 %	3.5 %	3.8 %
Completed ^e	2.2	2.2	2.4
Attempted ^f	1.0	1.1	1.2
Threatened ^g	0.5	0.6	0.6
Sexual assault ⁱ	2.6 %	2.5 %	2.8 %
Completed ^e	2.2	2.1	2.3
Attempted ^f	0.6	0.6	0.7
Threatened ^g	--	--	--

Note: Estimates using household or person base weights are the unweighted average of the five CBSA results.

-- Less than 0.05%

^aReflects the reciprocal of the probability of selection of the housing unit or college dormitory room.

^bReflects the adjustment of household-level base weights for subsampling a single respondent among eligible respondents.

^cReflects equalization of weights by multiplying them by a single factor in each of the CBSAs in order to treat each CBSA as if it were the same size as the others.

^dIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^eIncludes incidents where the offender completed the behavior.

^fIncludes incidents where the offender physically tried, but did not complete the behavior.

^gIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^hIncludes penetrative sexual contact using force or while unable to consent.

ⁱIncludes non-penetrative sexual contact using force or while unable to consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 10-2 compares CATI prevalence rates for rape and sexual assault with three sets of intermediate weights: the initial base weight as the inverse of the probability of selection of the phone number, the adjusted person-level weights just before compositing, and the composite weights. As with ACASI, the estimates in the table based on intermediate weights are the unweighted average of CBSA-level rates. Because of the overlap of the cell phone and landline universes, the initial base weights from the combined cell and landline samples do not properly account for the target universe, but the composite weights theoretically do so. Differences among the estimates in table 10-2 are modest. In particular, results based on the final weights are almost indistinguishable from those based on the composite weights.

In summary, the final weights for the ACASI sample include a raking step that compensates for the underrepresentation of some high-risk groups, particularly the youngest eligible respondents, resulting in a modest increase in key estimates. The CATI results appear quite insensitive to the choice of weight, even though only the composite and final weights are theoretically justified.

Table 10-2. Primary prevalence rates from CATI samples using intermediate and final weights, 2014–2015

Characteristic	Base weights ^a	Adjusted person weights ^b	Composite weights ^c	Final weights ^d
Rape and sexual assault ^e	5.6 %	5.3 %	5.3 %	5.3 %
Completed ^f	4.4	4.1	4.2	4.2
Attempted ^g	1.6	1.5	1.4	1.5
Threatened ^h	0.3	0.4	0.4	0.4
Rape ⁱ	3.2 %	3.2 %	3.1 %	3.1 %
Completed ^f	2.1	2.1	2.0	2.1
Attempted ^g	1.1	1.1	1.0	1.0
Threatened ^h	0.3	0.4	0.4	0.4
Sexual assault ⁱ	2.9 %	2.6 %	2.7 %	2.7 %
Completed ^f	2.5	2.3	2.4	2.4
Attempted ^g	0.5	0.4	0.5	0.5
Threatened ^h	--	--	--	--

Note: Estimates using base weights or composite weights are the unweighted average of the five CBSA results.

-- Less than 0.05%

^aReflects the inverse of the sampling probability for the telephone number, adjusted to remove telephone numbers found to be non-residential and to compensate for units with unknown residency status.

^bReflects within-household subsampling when the screener identified more than one eligible respondent, adjusted for screener non-response.

^cReflects compositing person-level weights by a weighting factor of 0.5 for respondents represented in both frames, that is, when they had a cell phone for their own use but could also be reached by landline in their household.

^dReflects equalization of weights by multiplying them by a single factor in each of the CBSAs in order to treat each CBSA as if it were the same size as the others.

^eIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^fIncludes incidents where the offender completed the behavior.

^gIncludes incidents where the offender physically tried, but did not complete the behavior.

^hIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

ⁱIncludes penetrative sexual contact using force or while unable to consent.

^jIncludes non-penetrative sexual contact using force or while unable to consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

10.2 Neighborhood Characteristics of Respondents and Nonrespondents

For the ACASI sample, the census tract was known for each sample address. In turn, the individual sample records were matched to tract-level characteristics from the 2014 five-year estimates from the ACS. The tract-level characteristics used were all expressed as percentages, such as the percent Hispanic. Within each CBSA, weighted averages were computed for the tract-level percentages for responding and nonresponding households, using the housing unit base weights. Tables 10-3 and 10-4 present the results for each CBSA and their average.

With few exceptions, the percentages for the neighborhood characteristics for respondents and nonrespondents are similar. For example, in all the CBSAs respondents lived in neighborhoods (that is, census tracts) that, on average, were composed of 34.7 percent Hispanic residents, while nonrespondents lived in neighborhoods composed of 33.3 percent Hispanic residents. One of the most prominent sets of systematic differences in table 10-3 is in the proportion of renters. A possible explanation for this difference is that interviewers experienced additional obstacles to contacting sample addresses if they were in neighborhoods with a high proportion of rental housing units. In other respects, however, the respondents in the sample were associated with slightly lower socioeconomic status in terms of education, income under \$50,000, and food stamp recipiency than nonrespondents. It is possible that the incentives were a stronger inducement to response for households in neighborhoods with these characteristics.

Table 10-3. Average percentages of tract-level characteristics of respondents (R) and nonrespondents (NR) for the ACASI sample, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015

Neighborhood characteristic	Overall		Phoenix		Los Angeles	
	R ^a	NR ^b	R ^a	NR ^b	R ^a	NR ^b
Hispanic	34.7 %	33.3 %	36.7 %	35.4 %	46.5 %	46.0 %
White only, non-Hispanic	42.7	44.5	50.7	51.8	32.7	31.3
Female	51.2	51.1	50.3	50.1	50.8	50.8
High school education or less	43.1	41.9	43.2	41.9	44.1	43.7
Uninsured	19.8	18.6	20.0	18.9	21.2	20.3
Unemployed	10.2	9.9	9.4	10.0	11.2	10.5
Income < \$50k	46.0	42.3	52.2	48.9	43.8	40.9
Below poverty	9.1	8.9	10.7	11.1	8.4	8.0
Receiving food stamps	13.3	12.1	15.2	14.8	8.5	8.4
Renter-occupied housing units	53.8	59.1	55.5	60.6	43.3	45.0
In same house 1 year ago	84.3	84.8	79.5	80.5	85.1	84.4
Non-family households	32.5	29.7	35.1	32.6	33.9	31.4
Never-married persons	37.5	35.3	37.7	34.6	41.0	39.8

^aIndicates respondents to the audio computer-assisted self interview.

^bIndicates nonrespondents to the audio computer-assisted self interview.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

Table 10-4. Average percentages of tract-level characteristics of respondents (R) and nonrespondents (NR) for the ACASI sample, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015

Neighborhood characteristic	Miami		New York City		Dallas	
	R ^a	NR ^b	R ^a	NR ^b	R ^a	NR ^b
Hispanic	35.9 %	37.0 %	27.5 %	23.5 %	26.9 %	24.6 %
White only, non-Hispanic	33.5	29.9	47.9	54.3	48.7	55.2
Female	52.0	52.2	51.9	51.3	50.9	51.3
High school education or less	44.2	43.2	43.3	42.2	40.9	38.6
Uninsured	24.6	24.0	12.2	11.2	21.0	18.7
Unemployed	12.2	12.0	10.2	9.8	8.2	7.4
Income < \$50k	49.9	48.0	39.7	35.5	44.3	38.3
Below poverty	9.8	9.7	8.0	7.5	8.5	8.1
Receiving food stamps	16.5	16.3	15.0	12.0	11.3	9.0
Renter-occupied housing units	61.1	65.1	52.1	59.8	57.1	65.0
In same house 1 year ago	83.8	83.6	90.5	92.2	82.4	83.4
Non-family households	31.2	28.8	31.9	28.9	30.2	26.7
Never-married persons	36.0	35.0	39.2	36.1	33.6	31.0

^aIndicates respondents to the audio computer-assisted self-interview.

^bIndicates nonrespondents to the audio computer-assisted self-interview.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

Tables 10-5 and 10-6 contain the county-level characteristics of CATI respondents and nonrespondents. The results show that the county-level characteristics for the respondents are strikingly similar to those of nonrespondents. This is not only true overall but for each of the five CBSAs as well. Unlike the ACASI results, the proportion of renters is almost perfectly balanced between respondents and nonrespondents. The socioeconomic characteristics are also in close alignment. As noted earlier, however, the restriction of the analysis to county-level characteristics may reduce the power to reflect differences for the CATI sample compared to the fine-grained analysis possible with the ACASI sample.

Table 10-5. Average percentages of county-level characteristics of respondents (R) and nonrespondents (NR) for the CATI sample, overall and for Phoenix and Los Angeles core-based statistical areas, 2014–2015

Neighborhood characteristic	Overall		Phoenix		Los Angeles	
	R ^a	NR ^b	R ^a	NR ^b	R ^a	NR ^b
Hispanic	33.8 %	34.5 %	29.9 %	29.9 %	44.6 %	44.8 %
White only, non-Hispanic	43.6	42.9	57.8	57.8	31.1	30.9
Female	51.1	51.1	50.5	50.5	50.7	50.7
High school education or less	40.2	40.4	37.0	36.9	41.2	41.4
Uninsured	18.9	19.0	16.4	16.4	19.7	19.7
Unemployed	9.7	9.7	9.4	9.3	10.5	10.6
Income < \$50k	44.7	44.9	46.7	46.7	42.3	42.5
Below poverty	9.6	9.6	10.7	10.7	8.7	8.8
Receiving food stamps	12.2	12.4	12.1	12.1	7.8	7.8
Renter-occupied housing units	56.6	56.1	61.9	61.7	49.3	49.2
In same house 1 year ago	85.0	85.0	80.7	80.7	86.3	86.4
Non-family households	33.5	33.6	34.0	34.1	31.7	31.8
Never-married persons	35.6	35.8	33.2	33.3	39.5	39.6

^aIndicates respondents to the computer-assisted telephone interview.

^bIndicates nonrespondents to the computer-assisted telephone interview.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

Table 10-6. Average percentages of county-level characteristics of respondents (R) and nonrespondents (NR) for the CATI sample, for Miami, New York City, and Dallas core-based statistical areas, 2014–2015

Neighborhood characteristic	Miami		New York City		Dallas	
	R ^a	NR ^b	R ^a	NR ^b	R ^a	NR ^b
Hispanic	40.6 %	43.4 %	23.2 %	23.3 %	30.6 %	31.2 %
White only, non-Hispanic	34.9	32.8	49.3	48.7	45.1	44.2
Female	51.5	51.5	51.9	51.9	50.7	50.7
High school education or less	43.2	43.9	38.9	38.7	40.9	41.1
Uninsured	23.8	24.3	11.9	12.0	22.5	22.7
Unemployed	11.3	11.3	9.3	9.3	8.1	8.2
Income < \$50k	51.3	51.8	38.3	38.3	44.8	45.1
Below poverty	10.4	10.5	9.0	9.1	9.1	9.1
Receiving food stamps	16.5	17.4	12.8	12.8	11.8	12.0
Renter-occupied housing units	62.0	61.2	51.0	50.2	58.6	58.2
In same house 1 year ago	85.1	85.4	89.7	89.5	83.1	83.0
Non-family households	35.0	34.7	35.3	35.7	31.6	31.8
Never-married persons	33.9	34.2	38.7	39.0	32.6	32.8

^aIndicates respondents to the computer-assisted telephone interview.

^bIndicates nonrespondents to the computer-assisted telephone interview.

Sources: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015, and United States Census Bureau, American Community Survey public file, 2014.

10.3 Geographic Precision of the CATI Sample

The ACASI sample was drawn directly from the resident population of the RSA Pilot Test CBSAs, while the CATI sample could only approximate the same target population. The ACASI and CATI samples differed in their coverage of two groups of respondents, *inmovers* and *outmovers*. For purposes of discussion in this section, inmovers will refer to eligible respondents with area codes not included in the CATI frame. A typical example would be a respondent whose cell phone number was outside the range of area codes sampled for RSA Pilot Test. In contrast, outmovers had area codes included in the RSA Pilot Test sample frame but did not live in any of the RSA Pilot Test CBSAs. The typical example of an outmover is a person who purchased a cell phone within a RSA Pilot Test CBSA and did not change her phone number when moving. Potential respondents can also be outmovers simply by living outside the CBSA boundaries but having a number that was included in the RSA Pilot Test frame.

As described section 3.3, the CATI landline frame comprises 100-banks of phone numbers. Using past data, a bank was included in the frame if it was determined to be predominantly located in one of the RSA Pilot Test CBSA counties. Similarly, the cell phone frame was constructed from 1000-banks of numbers assigned to wireless service. In addition to an effort to direct the sample to the CBSA counties, the questionnaire also collected information on the probable location of the respondent's residence. For landline respondents, one of the first introductory questions asked, "In what county (do you live/is this household located)?" The telephone interviewers could select from a list of the RSA Pilot Test counties, but respondents living in counties not on the list were coded "Other," don't know, or refusal. The interview was terminated for landline respondents in the "Other" category.

Cell phone respondents were also asked the question on residence, but those falling into the "Other" category were retained in the sample, with the rationale that the investment in reaching the cell phone respondent justified the additional marginal cost to complete the interview. If the rest of the interview was completed, the respondent was included in the final estimates. It is possible, however, to assess the impact of this decision by comparing respondents indicating that they lived in one of the 42 counties to those classified as "Other."

In addition to the initial question on county of residence in the CATI interview, two other items collected during the RSA Pilot Test also provide clues about the effect of geographic imprecision on the analysis:

- At the end of completed CATI interviews, respondents were asked their names and addresses, including zip code, to receive their incentive compensation. About 81 percent of cell phone respondents age 18-49 completing the interview provided this information. The respondent's zip code gives a separate indication of whether she lived in one of the five CBSAs.
- The ACASI respondents were asked their cell phone area code if they had one. When compared against the area codes of phone numbers included in the CATI sample, the ACASI respondents' area codes indicated in most cases whether they would have been included in the CATI frame.

Two definitions of inmovers are considered for the ACASI data:

1. the area code of the cell phone number was outside the specified CBSA
2. the area code of the cell phone number was outside all of the RSA Pilot Test CBSAs.

Three definitions of outmovers are considered for the CATI cell phone sample:

1. The respondent did not live in the target CBSA according to the response to the county of residence question.
2. The respondent did not live in any of the RSA Pilot Test CBSAs according to the response to the county of residence question.
3. If the respondent reported a zip code, the zip code was outside the RSA Pilot Test CBSAs, but if the respondent did not report a zip code, the response to the county of residence was outside the RSA Pilot Test CBSAs.

Table 10-7 summarizes the observed percentages of inmovers and outmovers based on these definitions. The first row of the table reports the percentage of outmovers excluded from the landline sample out of all cases with completed screening, including both cases identifying an eligible respondent and those where no household member was eligible. Overall, about 8 percent of the landline sample reported that they were not in any of the RSA Pilot Test counties.

The next three lines of the table apply the preceding three definitions of outmovers to the cell population, with results ranging from a high of 21.2 percent outmovers for the first definition to 16.2 percent for the third, expressed as proportions of the completed cell phone interviews. Because

migrants between the RSA Pilot Test CBSAs were still in the target population, the third and fourth rows more closely summarize the effect on the RSA Pilot Test of including respondents who lived outside of the target geographic area at the time of the survey. The fourth row implementing the third definition has the advantage of using all of the available information on respondents' probable residence. This approach is used to define cell phone outmovers in tables 10-8 and 10-9.

Overall, approximately 93.6 percent of respondents completing the ACASI interview reported having a cell phone and provided their area code. Of these, approximately 13 percent reported cell phone area codes outside their own CBSA, and approximately 11 percent reported area codes outside the entire set of RSA Pilot Test area codes, as shown in the fifth and sixth rows of table 10-7. The analysis is an approximation, because the CATI cell phone frame was defined by 1000 blocks of numbers rather than area codes. Because some of the area codes include areas in counties outside the RSA Pilot Test CBSAs, some of the cell phone numbers of ACASI cases may have been counted as included the CATI cell phone frame when they were actually inmovers. Thus, the two rows might underestimate the proportion of inmovers in the ACASI sample.

Table 10-7. Estimates of the proportion of outmovers and inmovers for ACASI and CATI samples, overall and by core-based statistical area, 2014–2015

	Overall	Phoenix	Los Angeles	Miami	New York City	Dallas
Outmovers ^a excluded from landline sample, not in RSA Pilot Test core-based statistical areas	8.3 %	7.2 %	9.4 %	10.0 %	9.3 %	5.7 %
CATI ^b cell outmovers ^a in the cell phone sample, not in same core-based statistical area	21.2	19.5	21.7	20.6	26.1	18.9
CATI ^b cell outmovers ^a in the cell phone sample, not in RSA Pilot Test core-based statistical areas	19.1	17.2	19.7	19.1	22.6	17.3
CATI ^b cell outmovers ^a in the cell phone sample, not in RSA Pilot Test core-based statistical areas according to the zip code provided	16.2	15.3	17.4	17.6	16.0	15.2
ACASI ^c cell inmovers ^d in the ACASI sample whose cell phones have area codes outside their core-based statistical area	13.0	17.9	8.1	10.8	10.6	15.3
ACASI ^c cell inmovers ^d in the ACASI sample whose cell phones have area codes outside RSA Pilot Test core-based statistical areas	10.9	15.0	5.8	8.7	8.6	13.9

Note: Estimates are based on weighted data.

^aRefers to respondents with phone numbers included in the RSA Pilot Test CATI sample frame, but did not live in any of the RSA Pilot Test CBSAs.

^bComputer-assisted telephone interview.

^cAudio computer-assisted self-interview.

^dRefers to respondents who live in one of the RSA Pilot Test CBSAs, but who have phone numbers not included in the RSA Pilot Test CATI frame.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

To illustrate the possible impact of the outmovers and inmovers on the comparison of the CATI and ACASI results, table 10-8 presents the age distribution of the CATI cell phone outmovers using the third definition, which incorporates information from respondents' zip codes when available; that is, the definition used in the fourth line of table 10-7. Similarly, the age distribution of the ACASI cell-phone inmovers is also given in table 10-8, using the second of the two definitions, shown as the sixth line of table 10-7. Outmovers and inmovers were disproportionately concentrated in ages 25-29.

Table 10-8. Age distributions of the CATI cell phone outmovers and ACASI cell phone inmovers compared to cell phone nonmovers, 2014–2015

Age	CATI ^a cell outmovers ^b	CATI ^a cell nonmovers ^c	ACASI ^d cell inmovers ^e	ACASI ^d cell nonmovers ^c
18-20	7.3 %	7.9 %	5.4 %	8.8 %
21-24	15.9	13.6	15.4	11.7
25-29	23.1	16.8	27.9	15.6
30-39	33.3	32.5	31.8	31.5
40-49	20.4	29.2	19.6	32.3

Note: Estimates are based on weighted data.

^aComputer-assisted telephone interview.

^bRefers to respondents with phone numbers included in the RSA Pilot Test CATI sample frame, but did not live in any of the RSA Pilot Test CBSAs.

^cRefers to respondents who live in one of the RSA Pilot Test CBSAs, but who have phone numbers not included in the RSA Pilot Test CATI frame.

^dAudio computer-assisted self-interview.

^eRefers to respondents with phone numbers included in the RSA Pilot Test CATI sample frame, who also live in one of the RSA Pilot Test CBSAs.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 10-9 compares estimated prevalence rates for key variables for the same distinctions between outmovers and inmovers as in table 10-8, namely, the third definition of CATI cell phone outmovers and the second definition of ACASI cell phone inmovers. Because the sample sizes for outmover and inmover estimates are small, the comparisons of estimates for the CATI cell outmovers compared to the CATI cell nonmovers are not statistically different from one another. The prevalence of rape and sexual assault and of rape differs significantly for ACASI cell inmovers compared to nonmovers and for ACASI cell inmovers compared to CATI cell outmovers. Confidence intervals on the estimated differences are wide, however. For example, the estimated 6.1 percentage point difference between ACASI cell inmovers and CATI cell outmovers has a corresponding approximate 95 percent confidence interval of (0.4, 11.8) percentage points, spanning a negligible to substantial difference.

Table 10-9. Prevalence rates of rape and sexual assault for CATI cell phone outmovers and ACASI cell phone inmovers compared to cell phone nonmovers, 2014–2015

	CATI ^a cell outmovers ^b		CATI ^a cell nonmovers ^c		ACASI ^d cell inmovers ^e		ACASI ^d cell nonmovers ^c	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Rape and sexual assault ^f	6.6 %	1.3 %	5.4 %	0.5 %	12.7 %	2.6 %	5.4 %	0.5 %
Rape ^g	2.9	0.9	3.1	0.4	9.0	2.5	3.3	0.4
Sexual assault ^h	4.3	1.1	2.8	0.4	4.6	1.2	2.6	0.4

Note: Estimates are based on weighted data.

^aComputer-assisted telephone interview.

^bRefers to respondents with phone numbers included in the RSA Pilot Test CATI sample frame, but did not live in any of the RSA Pilot Test CBSAs.

^cRefers to respondents with phone numbers included in the RSA Pilot Test CATI sample frame, who also live in one of the RSA Pilot Test CBSAs.

^dAudio computer-assisted self-interview.

^eRefers to respondents who live in one of the RSA Pilot Test CBSAs, but who have phone numbers not included in the RSA Pilot Test CATI frame.

^fIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^gIncludes penetrative sexual contact using force or while unable to consent.

^hIncludes non-penetrative sexual contact using force or while unable to consent.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The point estimates in table 10-9 suggest that coverage differences between the ACASI and CATI survey might affect the overall comparison:

- If cell phone inmovers differ from nonmovers by the amounts that the ACASI results suggest, then the CATI cell phone sample is missing a critical group that, if included, might raise the CATI estimates. For example, the overall rate for rape and sexual assault might increase by about 0.5 or 0.6 percentage points if cell phone inmovers could be added to the CATI sample. If the cell phone inmovers are excluded from the ACASI data, the overall rate for rape and sexual assault drops from 5.9 percent to 5.3 percent. The cell phone inmovers are a legitimate component of the ACASI estimate; however, the calculation illustrates the possible effect of excluding CATI inmovers from the CATI estimates.
- If the differences between cell phone outmovers and cell phone nonmovers are as small as the preliminary results suggest, then inclusion of cell phone outmovers in the CATI cell sample had little effect on the estimates. The inclusion of cell phone outmovers in the CATI estimates would not adequately compensate for the omission of cell phone inmovers.

Because the estimates for inmovers and outmovers have large standard errors, these conclusions are not firm, however. For example, within the confidence intervals, it is possible that prevalence rates for cell phone inmovers are smaller than they appear in table 10-9.

10.4 Level of Effort Analysis

The analyses at the beginning of the chapter did not find evidence of nonresponse bias for either the ACASI or CATI surveys. These analyses found that when comparing respondents and nonrespondents along demographic and economic characteristics, there were no significant differences. While very instructive, these analyses do not directly address whether particular outcomes of interest, such as the rape and sexual assault rate, exhibit bias due to nonresponse. For example, it may be that respondents and nonrespondents are similar in age but differ in their risk of sexual assault. An illustration of this is provided above when comparing inmovers and outmovers. These two groups have similar age distributions, but the inmovers have much higher rates of rape than the outmovers. This section assesses nonresponse bias by examining key measures of sexual victimization by the level of effort (LOE) that it took to interview the individual. This approach assumes that respondents who require a high level of effort to complete the survey are more similar to nonrespondents than those who require little effort to complete. Differences in the estimates of victimization by the LOE are interpreted as an indication of nonresponse bias.

For purposes of this chapter, LOE is measured by the number of times an interviewer had to try to contact the respondent to complete an interview. For both the CATI and the ACASI, this is the total number of times an attempt was made to complete both the household screener and the main interview. To conduct the analysis, two different LOE groups were formed. First, the distribution of the number of contact attempts was examined for completed general population interviews. Cases that were within the top 30th percentile of calls needed to complete were defined as “high-effort” cases. In the case of the CATI, the cutoff was at 7.1 calls. To form the group, all completed interviews requiring 8 or more calls were designated “high effort.” For the ACASI cases, the cutoff was at 3.8 visits to the household, meaning that all cases requiring 4 or more visits to the household were categorized as “high-effort” cases.

Next, both prevalence and incidence rates were estimated for rape, sexual assault, and other unwanted sexual contact within each of these two levels of effort groups. Independent t-tests were used to test whether the rates of sexual victimization differed significantly between the high and low effort groups. There were no statistically significant differences for either mode for any of the victimization rates between the two LOE groups. There was also no consistent pattern across the different types of victimization. For the ACASI cases, the prevalence rates were almost identical between the two groups (e.g., for rape, both are 3.8%). The incidence rate for rape for the higher effort group is a bit higher than the low effort group (60.0% vs. 44.6%), but the difference is not statistically significant (table 10-10). Similarly for the CATI cases, both the prevalence and incidence rates are similar when comparing the two LOE groups (table 10-11).

Table 10-10. Prevalence and incidence rates of sexual victimization by level of effort for ACASI respondents in the general population ages 18-49

	Level of effort	
	Low ^{a*}	High ^b
Prevalence per 100		
Rape ^c	3.8 %	3.8 %
Sexual assault ^d	2.9	2.7
Other unwanted sexual contact ^e	10.1	9.5
Incidence per 1,000		
Rape ^c	44.6	60.2
Sexual assault ^d	33.7	34.3
Other unwanted sexual contact ^e	97.7	99.6

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from low effort at the 95% confidence level.

^aLow-effort cases required 3 or fewer contacts to complete the household roster and main interview.

^bHigh-effort cases required more than 3 contacts to complete the household roster and main interview.

^cIncludes penetrative sexual contact using force or while unable to consent.

^dIncludes non-penetrative sexual contact using force or while unable to consent.

^eIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 10-11. Prevalence and incidence rates of sexual victimization by level of effort for CATI respondents in the general population ages 18-49

	Level of effort	
	Low ^{a*}	High ^b
Prevalence per 100		
Rape ^c	3.2 %	2.8 %
Sexual assault ^d	2.8	2.7
Other unwanted sexual contact ^e	5.6	7.1
Incidence per 1,000		
Rape ^c	45.4	38.1
Sexual assault ^d	34.1	33.5
Other unwanted sexual contact ^e	81.8	93.8

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from low effort at the 95% confidence level.

^aLow-effort cases required fewer than 8 calls to complete the screener and main interview.

^bHigh-effort cases required 8 or more calls to complete the screener and main interview.

^cIncludes penetrative sexual contact using force or while unable to consent.

^dIncludes non-penetrative sexual contact using force or while unable to consent.

^eIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The above analysis of LOE did not adjust for the variables included in the survey nonresponse adjustment. Even though the analysis did use the full set of survey weights, it is possible that the early responders differ along one or more of the other characteristics used in the nonresponse adjustment. Consequently, the above analysis may not fully reflect the full nonresponse adjustment implemented on the survey. For example, it might be the case that early responders are younger than late responders. Since younger people have higher rates of victimization, this would push the early group higher than the late group. By not accounting for this imbalance, the LOE analysis does not fully account for the fact that estimates in the body of the report adjusted for this characteristic.

To adjust for these variables, a series of logistic regressions predicting the same three victimization outcomes were estimated. The predictors in the regressions were the variables used in the nonresponse adjustment (age, race and Hispanic origin, marital status), as well as whether the survey was completed early or late. The early responders are coded as “0” and the late responders are coded as “1.” This means that a positive bias in the estimates is indicated by a regression coefficient of less than zero (i.e., the late responders have lower victimization rates or more favorable attitudes than early responders). The survey weights were used in these regressions and standard errors were estimated using the replication methods described in Chapter 3.

Consistent with the bivariate analysis reported above, the LOE variable was not statistically significant for most of the outcomes (table 10-12). The exception is for other unwanted sexual contact (non-RSA) in the CATI mode, which has significantly higher prevalence in the high-effort group than the low-effort group when accounting for respondent demographics. Assuming the late responders are more indicative of those who did not respond at all, this result suggests a negative bias in the estimates of other unwanted sexual contact for the CATI. The late responders have a higher rate than early responders, which means nonresponse tended to exclude those with higher rates. Level of effort was not a significant predictor for this victimization type in the ACASI mode, or of rape and sexual assault victimizations in either mode.

Table 10-12. Coefficients for logistic regression analysis predicting prevalence of rape, sexual assault and other unwanted sexual contact, using level of effort and respondent characteristics, for females ages 18-49 in the general population, 2014–2015

Parameter	Rape ^a		Sexual assault ^b		Other unwanted sexual contact ^c	
	ACASI ^d	CATI ^e	ACASI ^d	CATI ^e	ACASI ^d	CATI ^e
Intercept	-3.85 †	-3.86 †	-4.23 †	-4.22 †	-2.52 †	-3.09 †
Level of effort (LOE)						
Low LOE*						
High LOE	0.01	0.08	-0.03	0.06	-0.03	0.21 †
Age						
18-21*						
22-24	0.09	0.58	0.37	0.37	0.46 †	0.41
25-29	0.02	0.36	0.11	-0.11	0.36 †	0.00
30-34	0.13	-0.33	-0.43	0.13	-0.27	0.04
35-39	-0.81 †	-0.29	-0.28	-0.26	-0.53 †	-0.64 †
40-49	-0.24	-1.21 †	-0.51	-0.52	-0.94 †	-0.52 †
Race/Hispanic origin						
White only ^{f,*}						
Black only ^f	0.27	-0.07	0.11	0.05	0.03	-0.02
Hispanic	-0.46	-0.71 †	0.01	-0.48	-0.32 †	-0.64 †
Other ^{f,g}	-0.28	0.39	-0.78	-0.03	0.03	0.24
Education						
High School/GED*						
Less than high school	0.10	-0.04	-0.92 †	-1.04 †	-0.31	-0.55
Some college/associate's	-0.31	-0.03	0.26	0.67 †	-0.14	0.41 †
Bachelor's or more	-0.41	-0.07	0.04	0.39	0.30 †	0.31
Marital status						
Married*						
Not married	0.70 †	0.67 †	0.43 †	0.63 †	0.51 †	0.69 †

Note: Estimates are based on weighted data. See Appendix for standard errors.

* Comparison group.

† Significant difference from comparison group at the 95% confidence level.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes unwanted penetrative or non-penetrative sexual contact where force or while unable to consent was not reported while unable to consent

^dAudio computer-assisted self-interview.

^eComputer-assisted telephone interview.

^fExcludes persons of Hispanic or Latina origin.

^gOther race includes American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and persons identifying as "other" race (CATI only) or two or more races.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Overall, therefore, these LOE analyses are consistent with the other analyses of nonresponse bias discussed earlier in this chapter. There does not seem to be an indication of nonresponse bias for either the ACASI or CATI surveys for the rape and sexual assault estimates. Several other studies examining nonresponse bias for campus sexual assault surveys have found similar results (Cantor et al., 2016; Krebs et al., 2016). There are limitations to the LOE analysis—the major one being that it is not possible to verify the assumption that the late responders are representative of the nonresponders. Several studies have shown that this assumption does not always hold (Lin & Schaeffer, 1995; Olson, 2006). The response rates for this study limit the extent that nonrespondents can be characterized by late responders. The CATI survey interviewed 18 percent of the population, with late responders representing about one-third of these. It is hard to verify that these late responders represent the remaining 82 percent of those that did not respond. The ACASI survey had a somewhat higher response rate (40%), but it remains the case that late responders were used to represent the remaining 60 percent of the population.

11. Analysis of Vignettes, Event History Calendar, and Interviewer Behavior

This chapter describes several analyses that examine different aspects of the RSA Pilot Test instrument. The first section presents analysis of the vignettes presented at the end of the interview. The vignettes are intended to understand how respondents interpret several questions included on the RSA Pilot Test screening instrument and detailed incident form. The second section evaluates the effectiveness of the event history calendar. This was used on the ACASI instrument to promote recall and dating of sexual violence incidents. The third section reports on interviewer behavior during the CATI interview. This provides a measure of the extent to which interviewers were able to administer the RSA Pilot Test questions, as well as indications of questions that may have been problematic for respondents.

11.1 Analysis of Vignettes

Vignettes present fictional scenarios to which respondents are asked to react. After reading the vignette, respondents are asked the same questions included on the survey. Responses to the questions provide an indication of how respondents are interpreting the questions. This experiment was used as one of the ways to evaluate the validity of several key items on the RSA Pilot Test.

Vignettes have several advantages in the context of studying survey reports of rape and sexual assault (Martin, 2004). First, rape and sexual assault have relatively low incidence in the population, which makes it difficult to understand the response process in the context of a general population survey. The fictional scenarios used in vignettes are designed to be relevant to everyone in the population regardless of whether one has been victimized or not. Second, the use of fictional scenarios offers a less threatening way to ask respondents about the event (Finch, 1987). And third, vignettes allow manipulation of key aspects of the situation, such as use of alcohol and level of consent. This permits investigators to observe how respondents interpret questions across a number of different situations.

Vignettes have been used in previous studies to successfully understand perceptions of sexual consent and the role of contextual factors on perceptions of sexual assault (Humphreys, 2007; Lim & Roloff, 1999). In the RSA Pilot Test, each respondent was presented with two vignettes toward the end of the interview to help better understand how a diverse sample of women would

interpret different circumstances that affect reporting of rape. One vignette pertained to the role of alcohol and the other vignette pertained to the use of verbal pressure. The order of the vignettes was randomly assigned, as were the conditions within each vignette. A total of 48 different alcohol vignettes and 40 different verbal pressure vignettes were tested.

11.1.1 Alcohol Vignettes

The alcohol vignettes were administered to address four issues related to the interpretation of the RSA Pilot Test survey questions. The first research question is –

RQ1. How do respondents answer the alcohol-facilitated screening question (SV5) as the vignette describes increasing use of alcohol and increasing pressure to consume more alcohol?

“SV5. Has anyone made you have sex when you were unable to consent because you were too drunk, high or passed out?”

There are several possible ambiguities related to this screening question. One is how the question is interpreted when the subject is clearly impaired by alcohol, but is not unconscious. Another possible ambiguity is how the subject might respond if she was being encouraged to continue drinking alcohol in order to impair her physical and cognitive functioning.

The vignettes varied the extent to which a female consumed alcohol and the extent to which her male companion either knew she was getting drunk or actively encouraged her to get drunk (see methods section below).

RQ2. How do respondents answer the physical force screening question (SV1) as the vignette describes increasing pressure to consume more alcohol?

“SV1. Has a male used force or threats of force to make you have vaginal sex against your will? By vaginal sex, it means putting his penis in your vagina.”

Given alcohol's ability to inhibit physical functions, including the ability to resist force, varying levels of alcohol consumption may affect how respondents interpret the use of force in the screening item for rape by physical force (SV1). The analysis below assesses how respondents answer this question when the pressure to consume alcohol is varied.

RQ3. How do respondents answer questions about the details on alcohol facilitation?

The detailed incident form (DIF) asks a number of questions about the level of intoxication of the subject during the incident. One item of interest was whether or not the respondent thought the person was intentionally trying to get her drunk (RQ3.1).

“G7. Do you think he was trying to get you drunk so he could sexually take advantage of you?”

The vignettes vary the level of effort the male was using to try to get the subject drunk. The analysis below assesses how respondents answered this question based on those actions.

A second survey item of interest was the subject’s ability to provide consent given varying levels of alcohol consumption (RQ3.2).

G12a. “Did the alcohol make you unable to give consent?”

This question is used as part of the classification algorithm to determine whether the incident should be considered a rape when the victim was unable to consent. The vignette analysis explores how respondents interpreted this question, given varying levels of alcohol consumption.

RQ4. How do respondents’ answers to questions about alcohol facilitation vary by the relationship between the individuals, expressions of consent, and mode of interview?

The relationship between the two individuals is an important contextual variable on what is accepted as consent. Humphreys (2007), for example, found that explicit consent was more likely to be expected from individuals who had just started a relationship compared to those that have been in a long-term relationship. Similarly, perceptions of consent may be mediated by perceptions of the level of alcohol involvement. Lim and Roloff (1999) refer to this as “contextual discounting,” whereby judgments about consent are mediated by perceptions of impairment. One of their (unconfirmed) hypotheses is that impairment may lead to discounting explicit expressions of consent.

The analysis examines differences across mode of interview. Mode may impact responses in several ways. One may be related to whether or not respondents are more willing to report something (e.g., forced penetration; being unable to consent) to an interviewer. Since the vignettes are fictional, they are not likely to be as sensitive to the mode of interviewing. However, social

desirability may play a role. For example, recent publicity related to campus sexual assault may influence the answers respondents may provide for the role of alcohol in the scenario. Mode of interview may also be important because the presentation of the vignettes was slightly different in each mode. In the ACASI, the respondent read the vignette and was able to re-read it, as it was placed at the top of the ACASI screen. On the telephone, the interviewer read the vignette to the respondent. Unless the respondent asked for the vignette to be repeated, the respondent had to retain all of the critical details at that first reading.

Experimental Design

A total of 5,773 respondents from the general population sample, aged 18-49, were included in the analysis, including 2,721 interviews conducted in CATI and 3,052 interviews conducted in ACASI.

To examine the above research questions, three factors were varied in the alcohol vignette: drinking behavior, the relationship between the victim and the perpetrator, and the level of consent given by the victim. The structure of the vignette followed those used by Lim and Roloff (1999). The template for the vignette was –

Tom and Sue {INSERT RELATIONSHIP}. They both drank alcohol. {INSERT DRINKING BEHAVIOR}. They went back {home/to Tom's place}⁴⁵. {INSERT CONSENT}.

Respondents were randomly assigned to one of 48 conditions at the time of recruitment. Different text was inserted into the vignette based on the condition assignment (table 11-1).

⁴⁵Two different locations (home vs. Tom's place) were used in the alcohol vignette experiment. Location is not an experimental variable. For “first date” and “long-term relationship,” “home” was used as the location in the text, whereas “home” was used as the location if the relationship between the victim and the perpetrator was “married.”

Table 11-1. Experimental conditions for the alcohol vignette

Condition	Wording of vignette
Drinking behavior	
Neither drunk	<i>Both Sue and Tom ordered drinks, but neither became drunk.</i>
Sue gets drunk	<i>Both Sue and Tom ordered drinks. Sue clearly became very drunk.</i>
Sue gets drunk from social pressure	<i>Sue kept telling Tom she was not thinking straight and wanted to stop drinking. Sue kept on drinking because she wanted to be social and clearly became drunk.</i>
Sue gets drunk while Tom refills her glass	<i>Tom knew that Sue got drunk very easily but Tom kept filling Sue's wine glass. Sue clearly became drunk.</i>
Relationship	
First date	<i>Went on their first date.</i>
Long-term relationship	<i>Have been in a sexual relationship for three months and went on a date.</i>
Married	<i>Have been married for five years and went out for dinner.</i>
Consent	
Verbal consent	<i>Tom asked if she wanted to have sex. Sue said yes and they had sexual intercourse.</i>
Nonverbal consent	<i>Tom kissed Sue and they had sexual intercourse. Sue did not say anything at the time, but she did not want to have sex.</i>
Nonverbal nonconsent	<i>Tom kissed Sue. She tried to push Tom away, but did not actually say no. They then had sexual intercourse.</i>
Verbal nonconsent	<i>Tom kissed Sue. Sue said she did not want to have sex, but Tom ignored her and they then had sexual intercourse. Sue did not resist again because she was afraid of Tom.</i>

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The relationship includes three conditions representing an increase level of seriousness: first date (“*went on their first date*”), long-term relationship (“*have been in a sexual relationship for three months and went on a date*”), and married (“*have been married for five years and went out for dinner*”).

The drinking behavior includes four conditions that vary the extent to which the woman became drunk and the extent to which she was pressured to drink: neither drunk (“*Both Sue and Tom ordered drinks, but neither became drunk*”), Sue gets drunk (“*Both Sue and Tom ordered drinks. Sue clearly became very drunk*”), Sue drinking due to social pressure (“*Sue kept telling Tom she was not thinking straight and wanted to stop drinking. Sue kept on drinking because she wanted to be social and clearly became drunk*”), and Tom refilling Sue’s glass (“*Tom knew that Sue got drunk very easily but Tom kept filling Sue's wine glass. Sue clearly became drunk*”).

The level of consent includes four conditions that vary the extent to which consent was given and the type of consent: verbal consent (“*Tom asked if she wanted to have sex. Sue said yes and they had sexual intercourse*”), nonverbal consent (“*Tom kissed Sue and they had sexual intercourse. Sue did not say anything at the time, but she did not want to have sex*”), nonverbal nonconsent (“*Tom kissed Sue. She tried to push Tom away, but did not actually say no. They then had sexual intercourse*”), and verbal nonconsent (“*Tom kissed Sue. Sue said she did not want to have sex, but Tom ignored her and they then had sexual intercourse. Sue did not resist again because she was afraid of Tom*”).

One example is shown below:

“Tom and Sue went on their first date. They both drank alcohol. Both Sue and Tom ordered drinks, but neither became drunk. They then went back to Tom’s place. Tom asked if she wanted to have sex. Sue said yes and they had sexual intercourse.”

After answering the vignette, respondents answered two screener questions:

- SV1. “Has a male used force or threats of force to make Sue have vaginal sex against her will? By vaginal sex, it means putting his penis in her vagina.”*
- SV5. “Has anyone made Sue have sex when she was unable to give consent because she was too drunk, high, or passed out?”*

Next, respondents answered two alcohol-related questions from the detailed incident form:

- G7. “Do you think Tom was trying to get Sue drunk so he could sexually take advantage of her?”*
- G12a “Did the alcohol make Sue unable to give consent?”*

A detailed description of the analytic approach and findings appears in Appendix H. Below we summarize the major conclusions from the analysis.

Key Findings

The vignettes provide a measure of how respondents are interpreting selected questions on alcohol facilitation and physical force. The vignettes were designed to assess how respondents reported on situations that varied across several different contexts, including whether the individual was drunk, whether there was surreptitious administration of alcohol, the extent of (non)consent

and the relationship between the victim and the offender. The vignettes also tested whether there were any differences between ACASI and CATI interviewing modes. A number of analyses were carried out. These are summarized below. More detailed analyses, including regression analyses, are provided in Appendix H.

The first two research questions (RQ1 and RQ2) related to how respondents answer the alcohol-facilitated (SV5) and physical force (SV1) screening questions. Overall, the effects of the different contextual conditions are consistent with the intent of the items. However, there are some differences in how respondents interpret key concepts such as “unable to consent” and physical force. Some of these differences may be due to the specificity of the vignettes, but they also reflect differences in how respondents interpret some of the key concepts.

For the alcohol-facilitated rape item (SV5), the most significant predictor for whether a respondent said “yes” was whether or not the victim was drunk. There did not seem to be a large effect of the degree of surreptitious administration of alcohol, although the act of Tom refilling Sue’s glass did have the highest proportion saying “yes.” This suggests that respondents are placing more weight on the result that Sue was very drunk than on the intentions of the possible perpetrator. As a screening question in a two-stage design, this is a preferable result. The ultimate classification of the incident is based on the detailed incident form, which asks whether the individual was unable to consent, as well as other characteristics that may influence how the incident is classified (knowledge of surreptitious motives; signs of being drunk). It is preferable that the incident be screened in if the victim feels it qualifies.

When answering the alcohol-facilitated rape item (SV5), the relationship between the victim and perpetrator was significant, but had only a minimal effect on how respondents answered the items. The only significant effect was when the couple was married, but even here the percent that said “yes” when Sue was drunk was at least 70 percent and was as high as 85 percent for the nonverbal consent condition (table 11-2). This also indicates that respondents are interpreting the question as intended, by ignoring the victim-perpetrator relationship and concentrating on alcohol use and type of consent.

The type of consent had a significant effect. The critical condition was whether or not Sue gave verbal consent. When she did, respondents were much less likely to say she was made to have sex because she was unable to consent. Nonetheless, when verbal consent was described in the vignette, there was variation in how respondents interpreted the situation and responded to the question on alcohol/drug-facilitated events. When the victim gets drunk on her own and but gives

verbal consent, slightly less than half of the respondents (44%) said Sue was made to have sex because she was unable to provide consent (table 11-2, first date condition). While the percent goes up for the two other drinking conditions, the percentage is not much more than half of the respondents (e.g., 57%; 62%). Respondents seem to differ on whether the verbal consent was an indication of ability to consent. This ambiguity may be due to different interpretations of the circumstances of the vignette, which do not provide insights into exactly what Sue's state of mind is. However, it is also likely that respondents have different views on whether being drunk overrides verbal consent.

Table 11-2. Percentage reporting that someone made Sue have sex when she was unable to consent because she was too drunk, high, or passed out, by consent, relationship type and drinking behavior for females ages 18-49 in the general population, 2014–2015

Scenario	Drinking behavior			
	Neither drunk ^a	Sue gets drunk ^b	Sue drinking due to social pressure ^c	Tom refilling Sue's glass ^d
First date				
Consent				
Verbal consent ^e	3.6 %	44.4 %	57.5 %	62.1 %
Nonverbal consent ^f	21.9	85.2	77.4	86.5
Nonverbal nonconsent ^g	24.4	82.8	80.4	92.0
Verbal nonconsent ^h	25.7	88.1	87.4	89.4
Long-term relationship				
Consent				
Verbal consent ^e	3.9 %	46.8 %	50.8 %	65.6 %
Nonverbal consent ^f	22.0	72.2	76.8	84.8
Nonverbal nonconsent ^g	19.7	80.7	82.7	89.5
Verbal nonconsent ^h	26.2	86.3	86.7	86.3
Married				
Consent				
Verbal consent ^e	5.6 %!	38.0 %	40.0 %	55.0 %
Nonverbal consent ^f	17.1	71.4	70.6	84.9
Nonverbal nonconsent ^g	27.7	75.6	79.8	86.4
Verbal nonconsent ^h	27.6	86.7	82.5	92.6
Number of weighted sample cases	5,459,590	5,128,151	5,259,307	5,434,977

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes scenarios where both Tom and Sue ordered drinks, but neither became drunk.

^bIncludes scenarios where both Tom and Sue ordered drinks and Sue eventually became drunk.

^cIncludes scenarios where Sue kept telling Tom she was not thinking straight and wanted to stop drinking. Sue kept drinking because she wanted to be social, and clearly became drunk.

^dIncludes scenarios where Tom knew that Sue got drunk very easily and he kept refilling Sue's wine glass and Sue eventually became very drunk.

^eIncludes scenarios where Tom asked Sue to if she wanted to have sex. Sue said yes and they had sexual intercourse.

^fIncludes scenarios where Tom kissed Sue and they had sexual intercourse. Sue did not say anything at the time but she did not want to have sex.

^gIncludes scenarios where Tom kissed Sue. She tried to push Tom away, but did not actually say no. They then had sexual intercourse.

^hIncludes scenarios where Tom kissed Sue. Sue said she did not want to have sex, but Tom ignored her and they then had sexual intercourse.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Interpretation of the physical force screening item (SV1) provides several different views of what respondents are thinking when answering this item. For the scenario where neither Sue nor Tom is drunk, there is a significant jump in the proportion of “yes” responses when going from the consent to non-consent conditions. The percent that said “yes” in the nonverbal non-consent condition is around 40 percent (or slightly more) across the relationship conditions (table 11-3). There is a significant jump when there is verbal non-consent to around 70 percent or more. This is consistent with the intent of the question. Situations where verbal nonconsent is ignored and there was a threat of force should be considered a rape. On the other hand, a significant percentage of respondents (30%) did not think Sue was forced to have sex even though the scenario describes her as verbally refusing and she was afraid of what Tom might do if she resisted more. Some respondents may require more evidence of physical force to respond affirmatively to the force item.

Table 11-3. Percentage reporting that a male used force or threats of force to make Sue have vaginal sex against her will by consent, type of relationship and drinking behavior for females ages 18-49 in the general population, 2014–2015

Scenario	Drinking behavior			
	Neither drunk ^a	Sue gets drunk ^b	Sue drinking due to social pressure ^c	Tom refilling Sue's glass ^d
First date				
Consent				
Verbal consent ^e	3.2 %!	4.0 %	13.8 %	24.3 %
Nonverbal consent ^f	17.2	28.6	29.0	45.1
Nonverbal nonconsent ^g	41.4	52.6	49.6	68.5
Verbal nonconsent ^h	76.7	82.4	75.2	81.8
Long-term relationship				
Consent				
Verbal consent ^e	1.1 %!	6.5 %	12.9 %	20.1 %
Nonverbal consent ^f	17.7	26.9	31.5	42.6
Nonverbal nonconsent ^g	43.4	64.5	52.7	61.5
Verbal nonconsent ^h	80.9	76.5	74.8	79.3
Married				
Consent				
Verbal consent ^e	0.8 %!	5.0 %	11.4 %	14.8 %
Nonverbal consent ^f	12.0	36.3	33.4	46.5
Nonverbal nonconsent ^g	37.0	48.9	56.3	60.4
Verbal nonconsent ^h	71.7	86.9	86.2	80.6
Number of weighted sample cases	5,459,655	5,233,559	5,294,428	5,411,898

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes scenarios where both Tom and Sue ordered drinks, but neither became drunk.

^bIncludes scenarios where both Tom and Sue ordered drinks and Sue eventually became drunk.

^cIncludes scenarios where Sue kept telling Tom she was not thinking straight and wanted to stop drinking. Sue kept drinking because she wanted to be social, and clearly became drunk.

^dIncludes scenarios where Tom knew that Sue got drunk very easily and he kept refilling Sue's wine glass and Sue eventually became very drunk.

^eIncludes scenarios where Tom asked Sue to if she wanted to have sex. Sue said yes and they had sexual intercourse.

^fIncludes scenarios where Tom kissed Sue and they had sexual intercourse. Sue did not say anything at the time but she did not want to have sex.

^gIncludes scenarios where Tom kissed Sue. She tried to push Tom away, but did not actually say no. They then had sexual intercourse.

^hIncludes scenarios where Tom kissed Sue. Sue said she did not want to have sex, but Tom ignored her and they then had sexual intercourse.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

When there is excessive drinking involved (i.e., when both get drunk), perceptions of physical force are altered. There was both a main effect of drinking to saying “yes” to physical force, as well as an interaction with type of consent. If Sue was drunk, a larger percentage said there was force involved. There was a tendency to use the drinking condition as determining whether Sue was forced to have sex.

With respect to the question on surreptitious administration of alcohol (RQ3.1), the largest effect, as one might expect, was when Tom was pouring the drinks for Sue to get her drunk. However, there were a significant number of respondents (e.g., 40% – 50%) that thought Tom was trying to advantage of Sue when she was drunk, but there was nothing in the vignette to indicate that Tom was trying to take advantage of her (table 11-4). This is consistent with the results described above for the alcohol facilitation screening question (SV5). This may be due to the lack of specificity in the vignette with respect to Tom’s motives. But it reveals that a significant number of respondents view the end result of being drunk and assume it was facilitated by Tom.

Respondents’ interpretation of the DIF item on being able to give consent (RQ3.2) were very similar to the results for the alcohol facilitation screening question. The most significant condition was whether or not Sue was drunk. The extent to which there was pressure to get drunk did not seem to make a difference. This is consistent with the intent of the question, since it does not ask about surreptitious administration of alcohol. However, it re-emphasizes that many respondents are thinking being drunk is a sufficient condition to say that Sue was unable to consent, regardless of how much pressure there was for her to drink. The type of consent was also significant, with the critical condition being whether or not Sue gave verbal consent. However, like the results described above for the alcohol facilitation screening item (SV5), about half or more of the respondents believe that when Sue is drunk she is unable to consent, even when she gives explicit verbal consent (table 11-5).

Table 11-4. Percentage reporting that Tom was trying to get Sue drunk so he could sexually take advantage of her by consent, type of relationship and drinking behavior for females ages 18-49 in the general population, 2014–2015

Scenario	Drinking behavior			
	Neither drunk ^a	Sue gets drunk ^b	Sue drinking due to social pressure ^c	Tom refilling Sue's glass ^d
First date				
Consent				
Verbal consent ^e	7.1 %	38.6 %	59.5 %	89.2 %
Nonverbal consent ^f	20.6	58.5	71.9	94.8
Nonverbal nonconsent ^g	35.2	55.6	57.9	98.8
Verbal nonconsent ^h	48.0	69.3	74.7	97.0
Long-term relationship				
Consent				
Verbal consent ^e	5.9 %	27.3 %	45.0 %	78.1 %
Nonverbal consent ^f	27.6	42.5	55.5	82.9
Nonverbal nonconsent ^g	23.2	39.7	55.8	89.8
Verbal nonconsent ^h	45.6	49.2	59.1	90.9
Married				
Consent				
Verbal consent ^e	3.5 %	10.2 %	17.4 %	61.6 %
Nonverbal consent ^f	10.3	18.9	29.7	76.7
Nonverbal nonconsent ^g	20.7	20.8	29.5	81.5
Verbal nonconsent ^h	35.6	30.7	45.6	90.1
Number of weighted sample cases	5,321,955	4,922,626	4,959,766	5,289,509

Note: Estimates are based on weighted data. See Appendix A for standard errors.

[!] Interpret with caution. Coefficient of variation is greater than 50%.

^aIncludes scenarios where both Tom and Sue ordered drinks, but neither became drunk.

^bIncludes scenarios where both Tom and Sue ordered drinks and Sue eventually became drunk.

^cIncludes scenarios where Sue kept telling Tom she was not thinking straight and wanted to stop drinking. Sue kept drinking because she wanted to be social, and clearly became drunk.

^dIncludes scenarios where Tom knew that Sue got drunk very easily and he kept refilling Sue's wine glass and Sue eventually became very drunk.

^eIncludes scenarios where Tom asked Sue to if she wanted to have sex. Sue said yes and they had sexual intercourse.

^fIncludes scenarios where Tom kissed Sue and they had sexual intercourse. Sue did not say anything at the time but she did not want to have sex.

^gIncludes scenarios where Tom kissed Sue. She tried to push Tom away, but did not actually say no. They then had sexual intercourse.

^hIncludes scenarios where Tom kissed Sue. Sue said she did not want to have sex, but Tom ignored her and they then had sexual intercourse.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 11-5. Percentage reporting that alcohol made Sue unable to give consent by consent, relationship type and drinking behavior for females ages 18-49 in the general population, 2014-2015

Scenario	Drinking behavior			
	Neither drunk ^a	Sue gets drunk ^b	Sue drinking due to social pressure ^c	Tom refilling Sue's glass ^d
First date				
Consent				
Verbal consent ^e	10.9 %	52.3 %	68.0 %	70.2 %
Nonverbal consent ^f	35.2	90.6	86.1	93.3
Nonverbal nonconsent ^g	32.9	79.5	83.9	93.2
Verbal nonconsent ^h	25.5	69.4	80.3	77.5
Long-term relationship				
Consent				
Verbal consent ^e	7.1 %	45.4 %	64.0 %	70.0 %
Nonverbal consent ^f	30.4	84.0	84.6	76.5
Nonverbal nonconsent ^g	31.2	81.0	87.1	90.2
Verbal nonconsent ^h	24.4	73.1	77.1	69.8
Married				
Consent				
Verbal consent ^e	12.3 %	39.1 %	42.7 %	58.5 %
Nonverbal consent ^f	29.7	75.6	77.6	83.4
Nonverbal nonconsent ^g	24.7	80.2	89.7	85.8
Verbal nonconsent ^h	28.3	68.9	74.0	76.4
Number of weighted sample cases	5,403,542	5,147,576	5,195,805	5,378,671

Note: Estimates are based on weighted data. See Appendix A for standard errors.

^aIncludes scenarios where both Tom and Sue ordered drinks, but neither became drunk.

^bIncludes scenarios where both Tom and Sue ordered drinks and Sue eventually became drunk.

^cIncludes scenarios where Sue kept telling Tom she was not thinking straight and wanted to stop drinking. Sue kept drinking because she wanted to be social, and clearly became drunk.

^dIncludes scenarios where Tom knew that Sue got drunk very easily and he kept refilling Sue's wine glass and Sue eventually became very drunk.

^eIncludes scenarios where Tom asked Sue to if she wanted to have sex. Sue said yes and they had sexual intercourse.

^fIncludes scenarios where Tom kissed Sue and they had sexual intercourse. Sue did not say anything at the time but she did not want to have sex.

^gIncludes scenarios where Tom kissed Sue. She tried to push Tom away, but did not actually say no. They then had sexual intercourse.

^hIncludes scenarios where Tom kissed Sue. Sue said she did not want to have sex, but Tom ignored her and they then had sexual intercourse.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

There were very few interactions with the experimental variables. This suggests that the context of the situation, such as the relationship between the couple or the type of consent, did not have large effects on how respondents interpreted the questions (RQ4). The fact that relationship of the couple did not exert a large main effect and did not interact with any of the other variables (e.g., type of consent; drinking condition) indicates that respondents are interpreting these items as intended. Whether it was the first date, a long-term relationship, or the couple was married did not seem to affect whether respondents reported whether Sue was physically forced or was unable to consent. The one area that relationship did change responses was when asked about Tom's motives. Respondents were less likely to report a surreptitious motive when the couple in the vignette was in a long-term relationship and even less so when they are married.

Responses did not generally differ by interview mode. We observed a small number of inconsistent significant effects of mode across the questions that were evaluated in the vignettes. CATI respondents were significantly more likely to give positive responses with two items and less likely to give positive responses with one item. Even these statistically significant effects were substantively quite small. The patterns observed were the same across modes, even though respondents differed by how they were provided the vignettes (visual vs. verbally) and the extent to which the interview was private or confidential.

Older respondents, non-Hispanics, and whites were less likely to report incidents as rape. Education had different effects on the two types of questions. For the alcohol-facilitated item, higher education was associated with more reporting. The opposite was the case for the physical force item.

11.1.2 Non-Physical Pressure Vignettes

The second vignette was designed to assess how respondents answer questions on physical force involving vaginal penetration (SV1) and non-physical threats (SV6) to have sex. The intent of the initial screening question (SV1) is to capture situations where physical force was used. However, there are situations when respondents may interpret other types of non-physical pressure as being forced to have sex. RSA Pilot Test had a screening question asking about non-physical pressure to have sex (SV6):

SV6. Has anyone made you have any type of sex against your will by threatening to cause problems for you, such as at your job or school, at home, in your relationships, or in any other way?

The intent of this item was to capture incidents where some type of pressure was used that did not meet the legal criteria of a rape, but does involve coercive tactics. For example, threats to give the respondent bad grades or to cause problems at the job are coercive but may not meet a legal standard of rape. The vignette tested how respondents interpreted both SV6 and the physical force question (SV1) in situations that involved verbal pressure to have sex. Verbal pressure can come in several different forms (Livingston, 2007). The first is negative persuasion, where someone threatens a negative outcome (e.g., bad grades; to end the relationship). The second is positive verbal pressure, where the individual uses positive persuasion (e.g., take the relationship a step further). The third is a neutral form, typified by persistent attempts to have sex.

None of these forms of verbal pressure would qualify as a rape using the definition on the RSA Pilot Test, since they do not involve physical force or inability to consent. The research questions investigated how respondents answered the physical force and non-physical threat screening items for different forms of verbal pressure. The primary research question related to interpretation of the non-physical threat screening item (SV6).

RQ1: Does the rate of endorsement of SV6 vary directly with the type of verbal pressure used?

The vignette varied the extent to which negative, positive and neutral persuasion is described. It is expected that the vignette describing negative verbal persuasion will get the most endorsements of SV6 because it emphasizes this type of persuasion. The vignettes describing positive persuasion will get the fewest endorsements. The vignette describing neutral persuasion will fall in between the negative and positive conditions.

The second research question addresses how respondents answer the screening question on physical force (SV1) when the vignette describes only verbal pressure.

RQ2: Do respondents endorse SV1 when only verbal pressure is described in the vignette?

As noted in prior chapters, the screening questions include several different conditions related to endorsing this BSQ, including the tactic (physical force), behavior (vaginal penetration), and consent (against your will). Some respondents may not consider all three conditions when answering the question and endorse this item in response to vignettes describing negative verbal pressure.

Experimental Design

The structure of the vignette was –

“{INSERT RELATIONSHIP}. After dinner they went {to Mike’s home/home} to watch a movie together. During the movie, Mike began to kiss Becky {GENTLE FORCE}. He asked her if she would have sex with him. Beck said no, because she didn’t feel like having sex. {INSERT VERBAL PRESSURE}. They eventually had sexual intercourse.”

Three conditions of verbal pressure were tested (table 11-6). The negative condition was to end the couple’s relationship. The neutral condition was for the male to persistently ask the female for sex, and the positive condition was the male saying that sex would take the relationship in a positive direction. Like the first vignette, the situation also varied the relationship between the couple as a first date, in a sexual relationship, or married. A third set of conditions addressed whether or not “gentle force” was involved (Mike began removing Becky’s clothes).

An example of the vignette was –

“Mike and Becky went on a first date. After dinner they went to Mike’s home to watch a movie together. During the movie, Mike began to kiss Becky and he began to remove her clothes. He asked her if she would have sex with him. Becky said no, because she didn’t feel like having sex. Mike then said that he would not go out with her again if she did not have sex with him. They eventually had sexual intercourse.”

Table 11-6. Experimental conditions for the vignette involving verbal pressure

Condition	Wording of vignette
Type of verbal pressure	
Negative verbal pressure	
First date	Mike then said that he would not go out with her again if she did not have sex with him
Sexual relationship for 3 months, married	Mike then said he was going to end the romantic relationship if she did not have sex with him
Positive verbal pressure	Mike said "You're so beautiful. I really want to us to share something special"
Neutral verbal pressure	Mike also tried several times to persuade her to have sex
Relationship	
First date	Mike and Becky went on their first date
First date and work relationship	Mike is Becky's manager at work. They went on their first date
Sexual relationship for 3 months	Mike and Becky have been in a sexual relationship for 3 months and went on a date
Sexual relationship for 3 months and work relationship	Mike is Becky's manager at work. They have been in a sexual relationship for 3 months and went on a date
Married	Mike and Becky have been married for 5 years and went out for dinner
Gentle force	
No force	nothing
Use of gentle force	Mike began to remove her clothes

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Key Findings

The results associated with the first research question related to SV6 are consistent with expectations (table 11-7). Overall, the most significant variation in the rate of endorsement of SV6 is associated with the verbal pressure conditions. The level of those endorsing SV6 is, by a large margin, highest for the negative pressure condition. For example, 70 percent of those respondents that received the first date condition endorsed SV6. This goes up to 90 percent for those in the married condition. This compares to the positive verbal pressure, which varies between 5 percent to 25 percent, depending on the relationship. The neutral verbal pressure condition falls in the middle.

Table 11-7. Percentage reporting that a male used force or threats of force to make Becky have vaginal sex against her will by use of gentle force, type of relationship and coercion for females ages 18-49 in the general population, 2014-2015

Scenario	Type of coercion		
	End relationship if no sex ^a	Share something special ^b	Persuade to have sex ^c
First date			
Gentle force			
None	55.8 %	12.5 %	35.4 %
Mike began to remove Becky's clothes ^d	64.0	19.8	44.7
Work relationship ^e ; first date			
Gentle force			
None	69.8 %	24.1 %	42.2 %
Mike began to remove Becky's clothes ^d	64.9	30.3	40.7
Long-term relationship			
Gentle force			
None	67.3 %	13.3 %	46.9 %
Mike began to remove Becky's clothes ^d	69.7	18.5	43.2
Work relationship ^e ; long-term relationship			
Gentle force			
None	70.9 %	20.2 %	32.6 %
Mike began to remove Becky's clothes ^d	77.9	26.2	45.2
Married			
Gentle force			
None	72.8 %	11.6 %	37.3 %
Mike began to remove Becky's clothes ^d	75.8	30.0	40.8
Number of weighted sample cases	7,198,421	7,059,052	7,169,368

Note: Estimates are based on weighted data. See Appendix A for standard errors.

^aIncludes scenarios where Mike said he would not go out with Becky again or would end the romantic relationship if she did not have sex with him.

^bIncludes scenarios where Mike said "You're so beautiful. I really want us to share something special."

^cIncludes scenarios where Mike tried several times to persuade Becky to have sex.

^dIncludes scenarios manipulating whether or not Mike began to remove Becky's clothes during the fictional situation.

^eIncludes scenarios where Mike is Becky's manager at work.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

This pattern across the type of pressure is consistent with the intent of SV6, which is intended to capture the type of negative pressure described in the negative condition. Similarly, the relatively low percentage of individuals endorsing SV6 for the neutral and positive pressure condition also indicates the item was working as intended. There was a significant interaction between the type of relationship and verbal pressure. When the couple was married, respondents were more likely to endorse the item for the negative pressure condition. At least for this type of

negative pressure, therefore, the interpretation of this screener item is affected by the relationship between the couple.

Unlike the alcohol vignettes, there are differences in the rate of endorsement across the mode of interview. Those who responded on the ACASI were somewhat less likely to endorse the item than those on the telephone. This may be because this vignette was slightly longer than the first vignette, which may have made it harder to pick up all of the conditions when read over the telephone. Note that interviewers were trained to repeat the vignette if asked by the respondent.

With respect to the second research question, a significant number of respondents endorsed SV1 when there was negative pressure (table 11-8). For example, for the first date condition, 56 percent endorsed this item when there was no gentle force used and 64 percent when gentle force was used. This goes up to over 70 percent when the couple is described as married. This is not consistent with the intent of this item, which focuses on physical force. It may be that because Becky clearly says no and they have sex anyway, some respondents inferred that force was used. Nonetheless, this suggests that this item may elicit reports that do not fit the intended tactic. For example, respondents may have endorsed the item because they believe the event was something that qualified as being against Becky's will, although not something that necessarily involved physical force.

There was a large effect of education. Those with a high school education or less were more likely to endorse this item when compared to those with more education. Education is commonly used as a proxy for cognitive abilities. As noted above, this vignette was longer than the one on alcohol and may have been interpreted differently by respondents with different levels of education.

Table 11-8. Percentage reporting that someone made Becky have sex against her will by threatening to cause problems for her, by use of gentle force, type of relationship and coercion for females ages 18-49 in the general population, 2014–2015

Scenario	Type of coercion		
	End relationship if no sex ^a	Share something special ^b	Persuade to have sex ^c
First date			
Gentle force			
None	70.5 %	4.6 %	20.4 %
Mike began to remove Becky's clothes ^d	74.6	6.8	16.1
Work relationship ^e ; first date			
Gentle force			
None	77.8 %	15.6 %	39.2 %
Mike began to remove Becky's clothes ^d	75.2	24.6	25.4
Long-term relationship			
Gentle force			
None	74.2 %	3.9 %	14.7 %
Mike began to remove Becky's clothes ^d	84.9	8.4	20.1
Work relationship ^e ; long-term relationship			
Gentle force			
None	86.1 %	18.5 %	24.6 %
Mike began to remove Becky's clothes ^d	87.6	16.5	35.5
Married			
Gentle force			
None	89.7 %	5.9 %	15.9 %
Mike began to remove Becky's clothes ^d	87.9	8.1	24.8
Number of weighted sample cases	7,155,640	7,047,567	6,976,351

Note: Estimates are based on weighted data. See Appendix A for standard errors.

^aIncludes scenarios where Mike said he would not go out with Becky again or would end the romantic relationship if she did not have sex with him.

^bIncludes scenarios where Mike said "You're so beautiful. I really want us to share something special."

^cIncludes scenarios where Mike tried several times to persuade Becky to have sex.

^dIncludes scenarios manipulating whether or not Mike began to remove Becky's clothes during the fictional situation.

^eIncludes scenarios where Mike is Becky's manager at work.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

11.2 Event History Calendar

This section presents results on the use of the event history calendar (EHC) for the ACASI interviews. The EHC is a memory aid designed to improve the recall of retrospective events. The intent is to have the respondent use a calendar featuring personalized landmark events. The calendar displayed the reference period on the horizontal axis and personalized landmark events on the vertical axis. Providing landmark events over the 12-month reference period is intended to help the respondent focus on what happened over that period. One goal is to increase the ability to recall incidents. A second goal is to assist the respondents in dating incidents that are remembered. Of particular concern is reducing “external telescoping,” which is the tendency for respondents to remember events that occurred outside the reference period but misdate them as occurring within the reference period. As noted in section 9, there was evidence of external telescoping for the ACASI interviews.

11.2.1 Design and Administration

The EHC was administered using a two-step process. First, interviewers and respondents completed a paper-and-pencil calendar together. The interviewer protocol was flexible. There were specific questions that needed to be asked, but the order in which they were asked was dependent on the types of events that the respondent reported. The calendar listed the 12-month reference period along the top row and five event categories along the side, including family milestones (e.g., birthdays, wedding, births, or deaths), vacations, job changes, changes in residence, and other significant events (exhibit 11-1). An “X” was placed in a cell if the respondent experienced an event within a particular month.

Second, upon reviewing the calendar, the interviewer administered a set of prompts using computer-assisted personal interviewing (CAPI) to encourage respondents to report events in each of the five event categories on the calendar. A prompt, such as “Did anyone in your family take a vacation in the last 12 months?,” was administered only if the respondent did not already report an event in a specific category. Within the CAPI instrument, the interviewer then recorded if the respondent had entered at least one event in a given category (in either step of the process) or if they recorded zero events in a given category.

Exhibit 11-1. Event history calendar

Event Calendar Date ___ / ___ / ___ Case ID: 2 ___ - ___ - ___

PUT "X" IN CELL IF EVENT OCCURRED	12 MONTHS AGO	11 MONTHS AGO	10 MONTHS AGO	9 MONTHS AGO	8 MONTHS AGO	7 MONTHS AGO	6 MONTHS AGO	5 MONTHS AGO	4 MONTHS AGO	3 MONTHS AGO	2 MONTHS AGO	LAST MONTH	CURRENT MONTH
Birthday or birth of new child													
Change in marital status													
Death in family													
Vacation													
New job or promotion													
Moved to new house or apartment													
Other _____													

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Once the calendar was completed, it was placed next to the computer so the respondent could refer to it when filling out the rest of the questionnaire.

11.2.2 Events Reported on the Calendar

Overall, most respondents reported at least one type of event on the calendar (97.8% for the general population and 94.8% for the volunteer sample). Table 11-9 shows the percentage of respondents who entered an event in each category. Most respondents entered an event in the family milestone category. Specifically, 92.3 percent of general population (GP) respondents and 88.8 percent of volunteer (VO) respondents entered an event on the calendar for a birthday, birth of a new child, change in marital status, or death in the family in the last 12 months. In addition, 62.4 percent of GP respondents entered a vacation on the calendar, 38.9 percent entered a work-related event, 24.8 percent entered a change in residence, and 22.9 percent entered another type of event. Among the VO sample, 57.7 percent of respondents entered a vacation, 66.2 percent entered a

work-related event, 46 percent entered a change in residence, and 35 percent entered another type of event on the calendar.

Table 11-9. Type of event history calendar entry, by sample type, 2014–2015

	General population ^a	Volunteer sample ^b
Family milestones ^c	92.3 %	88.8 %
Vacation	62.4 %	57.7 %
Work-related events	38.9 %	66.2 %
Change in residence	24.8 %	46.0 %
Other events	22.9 %	35.0 %
Number of unweighted sample cases	3,053	984

Note: See Appendix A for standard errors. Event history calendar was administered only on ACASI.

^aFemales ages 18 to 49 in the general population, weighted data.

^bFemales ages 18 to 29 in the volunteer sample, unweighted data.

^cThis includes event history calendar entries for birthdays, birth of a new child, change in marital status, or death in the family.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

11.2.3 Respondents' Use of the EHC During the Interview

One measure of the effectiveness of the calendar was how often the respondent referred back to it during the interview. This is an indication that the EHC was useful in helping the respondent recall the specific timing of events that had happened in the past 12 months. Interviewers were instructed to make a note after the interview on whether the respondent referred to the EHC at any point during the interview. Interviewers could report that the respondent used the calendar, didn't use it, or they didn't notice it being used. Most respondents did not use the calendar during the interview (table 11-10; 78.4% for the general population and 74.7% for the volunteer sample). A slightly higher percentage of VO respondents (10.8%) were observed using the calendar than GP respondents (4.0%). A considerable number of interviewers didn't notice respondents using the calendar during the interview (17.6% of GP respondents and 14.5% of VO respondents).

Table 11-10. Respondents' usage of the event history calendar during the interview, by sample type, 2014–2015

	General population ^a	Volunteer sample ^b
Interviewer observed the respondent using the event history calendar?		
Yes	4.0 %	10.8 %
No	78.4 %	74.7 %
I did not notice	17.6 %	14.5 %
Number of unweighted sample cases	3,042	972

Note: See Appendix A for standard errors. Event history calendar was administered only on ACASI.

^aFemales ages 18 to 49 in the general population, weighted data.

^bFemales ages 18 to 29 in the volunteer sample, unweighted data.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Effect on Data Quality

The purpose of the calendar is to make the response task easier for the respondent. By going over what happened during the reference period, the respondent is asked to remember events or time periods that should help recall victimizations that might have occurred. One measure of whether this was successful is to see how the number of event types entered on the calendar affects response time, amount of missing data, and telescoping for the incident dating questions on the victimization screener.

A regression model was estimated for each of the outcomes (response time, the amount of missing data, and external telescoping). Each model contained variables for calendar completion and respondent characteristics, including age, race and ethnicity, education, income, and number of incidents reported. The amount of effort devoted to the calendar was operationalized by the number of events that were entered on the calendar. If zero or one event was entered, it was considered 'low effort'. If two to five events were entered, it was considered 'more effort'. Response time was a continuous outcome that represented the total time it took to complete the incident dating items. The amount of missing data and external telescoping were operationalized into binary variables. The respondent was considered to have missing data if they gave at least one "don't know" or "refused" response for the date of any incident. There was external telescoping if at least one incident date was reported outside of the 12-month reference period.

Table 11-11. Regression coefficient for number of events entered on the event history calendar for equations predicting response time, amount of missing data, and external telescoping in the Incident dating questions for females ages 18 to 49 in the general population, 2014–2015

Number of event types ^c	Response time ^a		Missing data ^b		External telescoping ^b	
	Coefficient	P-value	Odds ratio	P-value	Odds ratio	P-value
Zero to one	2.32	0.97	2.13	0.52	1.23	0.64
Two to five ^d	~	~	~	~	~	~
Number of unweighted sample cases used in the model	380		384		384	

Note: The event history calendar was only administered on ACASI.

~ Not applicable.

^aLinear regression that included controls for respondent demographic characteristics including race and ethnicity, age, income, education, and number of incidents reported.

^bLogistic regression that included controls for respondent demographic characteristics respondents race and ethnicity, age, income, and education.

^cMain independent variable in each regression model.

^dReference category for the variable “number of event types.”

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The regression coefficient and odds ratios for the effort variable are presented in table 11-12. The regression coefficient for completing the calendar is not statistically significant for any of the models, indicating that the calendar did not have an appreciable effect for any of the outcomes tested. The insignificant effects of the calendar on external telescoping are consistent with the results reported earlier in the report (see section 9), which found evidence that there was some external telescoping in the reports of rape. The insignificant regression coefficient for the telescoping equation further indicates that the number of events entered on the calendar was not highly correlated with external telescoping.

Respondents' Completion of the EHC and Reporting Victimization

Besides reducing external telescoping, a second way the calendar can improve data quality is by increasing the respondent's ability to recall an incident. By focusing the respondent on what was happening in her life over the reference period, she may be more likely to remember events. To investigate this possibility, the analysis examined the correlation between victimization reporting and use of the calendar.

Table 11-13 shows the association between the number of event types entered on the calendar and reporting a victimization in the last 12 months for the GP sample. The percentage of respondents reporting a past 12-month victimization seems to significantly increase as more event types are recorded on the calendar. The percentage of respondents reporting a past 12-month victimization progresses from 2.0 percent when zero event types are on the calendar to 30.6 percent when five event types are on the calendar. This suggests that respondents who put more effort into completing the calendar are more likely to remember and report a past 12-month victimization. The calendar could be helping these respondents effectively frame the 12-month reference period with personal landmarks, which improved the recall of retrospective victimizations.

The above relationship may be confounded by other factors influencing reports of past 12-month victimizations, such as the respondents' age, race, education, or income. To account for these characteristics, a logistic regression was estimated that predicted victimization based on the number of events reported and controlled for respondent demographic characteristics (table 11-14). The association between victimization and number of events on the EHC largely holds. The significant estimates show that respondents with three or four event types on the calendar were twice as likely to report a past 12-month victimization compared to respondents with one event type on the calendar. Respondents with five event types on the calendar were three times as likely to report a past 12-month victimization compared to respondents who reported one event type on the calendar.

Table 11-12. Victimization reports in the last 12 months, by number of event types entered on the event history calendar for females ages 18 to 49 in the general population, 2014–2015

	Number of event types					
	0*	1	2	3	4	5
Victimization reported in the last 12 months						
Yes	2.0 %	7.6 %†	9.8 %†!	16.9 %†	19.5 %†	30.6 %†
No	98.0 %!	92.4 %†	90.2 %†!	83.1 %†	80.5 %†	69.4 %†
Number of weighted sample cases	246,829	1,954,521	4,067,438	3,189,833	1,538,610	283,064

Note: See Appendix A for standard errors. Event history calendar was administered only on ACASI. Number of event types ranges from 0 to 5. There are five different types of events that respondents are prompted to report on the calendar: family milestones, vacation or family events, work-related events, change in residence, and other.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 11-13. Regression results predicting a past 12-month victimization based on the number of event types entered on the event history calendar for females ages 18 to 49 in the general population, 2014–2015

Number of event types	Odds ratio	P-value
0	0.2	0.39
1 ^a	~	~
2	1.3	0.22
3	2.1	0.001
4	2.2	0.001
5	3.3	0.001

Note: The event history calendar was administered only on ACASI. This table shows the results of a logistic regression that included controls for respondent demographic characteristics including race and ethnicity, age, income, and education. The model was based on 2,747 cases.

~ Not applicable.

^aReference category for the variable “number of event types.”

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

For the volunteer sample, there is a curvilinear relationship between reporting a past 12-month victimization and number of event types (data not shown). The logistic regression indicated that those with zero or three event types on the calendar were twice as likely to report a past 12-month victimization compared to those with one event type on the calendar. Those with four or five event types on the calendar were three times as likely to report a past 12-month victimization compared to those with one event type.

It isn't clear why the VO and GP results are slightly different. The higher victimization rates for those in the VO sample that did not report any events on the calendar are based on a relatively small number of respondents (51). Consequently, it may be that those who didn't report any EHC events on the VO sample represent an anomaly. The VO sample consists of a highly selective group of individuals who volunteered for the study rather than being part of the probability sample. Notwithstanding this inconsistency, the positive relationship between the number of events reported and victimization does suggest that the EHC may have helped recall. One cannot rule out an alternative explanation: that the use of the EHC is a measure of respondent motivation to fully participate in the interview. Those reporting fewer events on the calendar are less motivated than those who report many events. This motivation may also be reflected in the number of victimizations that are reported on the survey. It is not possible to parse out this “motivation” hypothesis from the recall hypothesis in the current study.

11.3 Interviewer Performance

The CATI version of the RSA Pilot Test involves interviewers administering highly sensitive and explicit questions. In contrast to a self-administered mode, like the ACASI, the interviewer has to implement procedures to deliver the survey questions. This section examines two different aspects related to the interviewers. One is the extent to which they were able to administer the questions as intended by the survey designers. This analysis uses behavior coding of key portions of the instrument. The second analysis conducts an evaluation of interviewer variance for key analytic outcomes (e.g., rape and sexual assault).

11.3.1 Behavior Coding

This section presents highlights of observations from behavior coding telephone interviews of 125 general population (GP) and 75 volunteer sample⁴⁶ (VO) respondents. Behavior coding is one method to evaluate the performance of survey items on a survey that is administered by an interviewer. Behavior coding was conducted to detect instances of social and cognitive difficulty within particular sections of the questionnaire. The method consists of listening to the interaction between the respondent and interviewer and coding what happens when a particular question is administered. This includes coding both the interviewer behavior (e.g., was the question worded as written? If not, was there a major change in the meaning?) and respondent behavior (did the respondent ask for clarification?; was an adequate answer given?). Items are evaluated by the extent to which the interaction deviates from what is expected.

The behavior coding targeted several different portions of the RSA Pilot Test questionnaire:

1. The informed consent statement. There are a number of conditions related to the consent statement, which may make it difficult for the interviewer to read and for the respondent to understand.
2. Demographic items and activity items. These items are relatively straightforward, with the possible exception of the items on activities.

⁴⁶This includes four cases with partial audio recordings that were kept in the data set because most sections had been coded.

3. Sexual victimization screening items. These items are very sensitive. They may also be difficult to answer because they contain several different concepts (e.g., behaviors; tactics; non-consent).
4. Behaviors and tactics on the detailed incident form. These items were used to classify the events into crime categories. Given their prominence in the classification scheme, it was important to evaluate their performance.
5. Vignettes. These items involve reading a short vignette (“story”) to the respondent and then asking questions about it. There was some concern whether respondents would be able to be able to carry out this task over the telephone.

While there have been a number of general population surveys that have contained similar content, there has been very little data published on how well interviewers are able to administer them. Of particular interest was how well the interaction went for the screening items, which introduce topics and use language that is not generally covered on a survey administered to the general population. Overall, the complete set of behavior coding results can be found in Appendix H, organized by section of the interview and sample type.

Design

The sample of cases selected for behavior coding were stratified by sample type and type of incident reported:

- stratum 1: No sexual victimization reported, n=39
- stratum 2: Past 12-month incidents involving unwanted penetration acts, n=58
- stratum 3: Past 12-month incidents involving unwanted sexual contact, no past 12 month incidents of unwanted penetration, n=55
- stratum 4: Lifetime incidents only, no past 12-month incidents, n=48.

A coding scheme measuring the initial reading of the question, the initial answer, and respondent and interviewer follow-ups was developed. A brief description of the coding scheme is presented below but it can be found in its entirety in Appendix H.

The interviewer's initial reading of the question was described using five codes:

- the question was read exactly as worded
- there was a major change to the question wording such that the question's meaning is altered
- the respondent interrupted the reading of the question
- the interviewer confirmed information from the respondent without reading the question
- the interviewer incorrectly skipped a question without reading it.

Coders were able to use a combination of these codes to describe the interviewers reading of the question when necessary (e.g., if the respondent interrupted the interviewer's reading of the question to provide an answer but the interviewer continues to read the question verbatim, this is coded as read the question "exactly as worded" and "respondent interruption").

The respondent's initial answer was described using five codes:

- adequate answer (i.e., answer fits one of the provided response categories)
- inadequate answer (i.e., answer doesn't fit one of the provided response options)
- qualified answer (i.e., answer fits one of the response categories but indicates doubt or uncertainty on the part of the respondent like "I think the incident happened in May")
- don't know, or refused question. Coders could use only one code to describe the respondent's initial answer.

Interviewer follow-up consisted of six behavior codes:

- interviewer repeats the question
- provided inappropriate feedback (i.e., interviewer conveys personal opinions about the survey or respondents situation)
- confirmed an unclear answer
- correct probe

- inappropriate probe
- failed to probe.

Probes are follow-up questions or comments that interviewers may use to obtain an adequate answer from respondents during the question-answer process. A “correct” probe is nondirective and should not affect the respondents’ answer (e.g., if a respondent wants a term defined, an interviewer may just say “it’s whatever it means to you”). An inappropriate probe influences a respondent and may lead them to a specific response.

Respondent follow-up consisted of five behavior code:

- pause/fillers (i.e., respondents pauses or hesitates for more than 2 seconds during the question-answer process)
- request for clarification on the question
- request for the question to be repeated
- change answer from the initial response
- comment on the sensitivity of the survey content.

Coders could code more than one interviewer or respondent follow-up code (e.g., respondent requests clarification on question meaning, interviewer correctly probes by saying “It’s whatever it means to you,” interviewer repeats question, and respondent changes their answer).

Six coders behavior-coded the 200 CATI cases. Roughly 20 percent of the cases were coded by two individuals to evaluate the intercoder reliability. This resulted in five coder pairs completing the double-coding of cases. Reliability was measured by comparing the behavior codes used by each pair in their double-coded cases. Each coder pair had a Kappa of above .70.

Results – General Population

This section describes the results for each section of the questionnaire for the general population sample.

Consent Form. Interviewers administered consent statements about the topics covered in the interview, voluntary participation, confidentiality, the procedure for expressions of harm, and

explicit language used in the interview with little difficulty. Each consent statement was read exactly as worded by the interviewer over 98 percent of the time.

Demographics. Items about the respondent's activities, income, home ownership, and length of stay were behavior coded. Overall, interviewers read the question exactly as worded the majority of the time; however, a particularly low percentage of respondents answers were adequate for the income question (72.5%) and the question on how long the respondent has been at their address (82.1%). The respondents who didn't provide an adequate answer on the income question primarily stated that they didn't know their total household income (10.9%). When asked about how long they have lived at their address, many respondents provided qualified answers that indicated doubt such as "I think about five years" (11.7%).

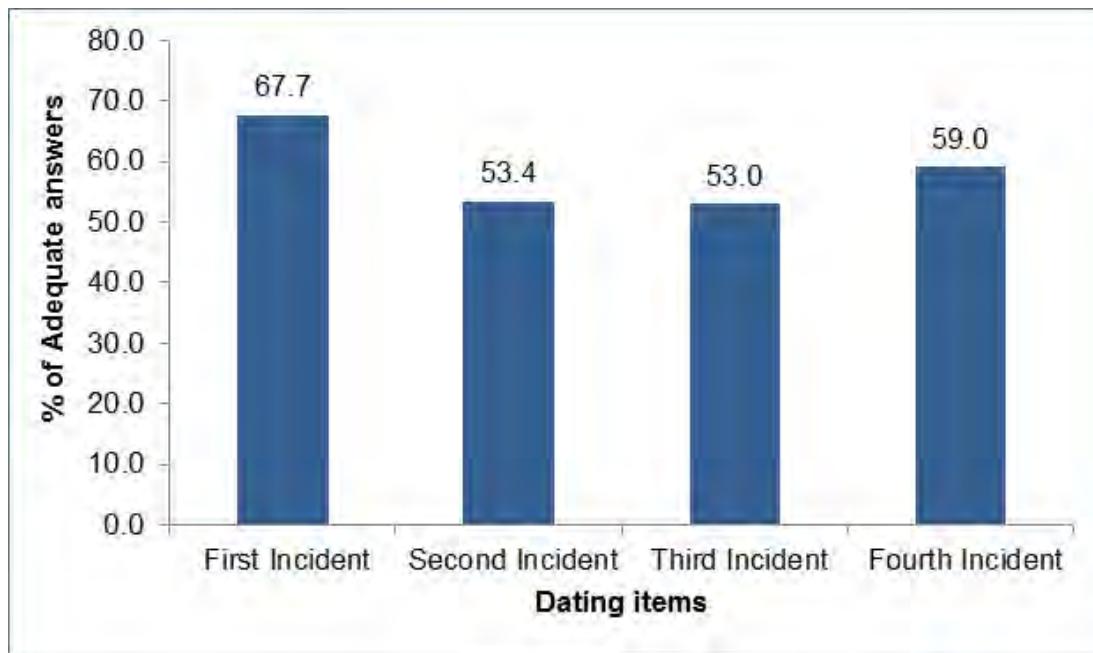
Sexual Victimization Screener. When designing the survey, one item of particular concern in the sexual victimization screener was the item attempting to de-duplicate incidents, which asked, "Is this part of the other incident you already reported in (month/year)?" Behavior coding showed that respondents had no trouble with this item, answering adequately 98.8 percent of the time with only 3.3 percent of respondents asking for the question to be repeated and no respondents asking for clarification.

Elsewhere in the sexual victimization screener, however, issues were identified from the behavior coding. Respondents interrupted the interviewer during the first two screener items, regarding vaginal sex and oral sex, about 11 percent of the time and interrupted the third screener item, regarding anal sex, 9 percent of the time. This could be related to the respondents being unaccustomed to the explicit language and sensitive subject matter at the beginning of the interview. Interruptions for subsequent screener items dropped to below 6 percent, possibly signaling respondents had gotten used to the questions or did not find the other items to be as sensitive. Furthermore, respondents asked for clarification on the first screening item (vaginal sex) 12.6 percent of the time, whereas the other screener items required clarification less than 5 percent of the time (e.g., "in my entire life?"; "like force you angrily?"). This could be related to respondents acclimatizing to the interview structure.

Respondents were asked to provide the month and year of up to four incidents of each type of sexual victimization that happened to them. A relatively high percentage of respondents did not initially provide an adequate answer. Respondents adequately dated the most recent incident 67.7 percent of the time and dated the second, third, and fourth most recent incidents adequately an average of 55 percent of the time (figure 11-1). Among the respondents who didn't provide an

adequate answer for the date, many said they didn't know the date, provided an inadequate answer like "it happened in October or November," or provided a qualified answer like "I think it happened in October." Respondents changed their answer to the date of the incident roughly 7 to 15 percent of the time.

Figure 11-1. Percent of adequate answers by victimization screener incident dating items for females 18+ in the general population (weighted)



Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Behavior and Tactics. In the detailed incident form (DIF), behavior coding revealed potential cognitive difficulties with several items. Items in the series D1 and D2 determine the unwanted behaviors associated with the incident, and whether they were threatened, attempted, or completed. For D1a, for example, the question was:

Did the person threaten to, try to or actually put his penis in your vagina when you didn't want it to happen?

Say one if the person verbally threatened to do this. Two if the person physically tried to do this. Three if the person actually did do this. Four if this did not happen.

This type of question was repeated for three other penetrative behaviors (items D1b – D1d) and five other types of sexual contacts (D2a – D2e). For the first few items on penetrative behaviors, a number of respondents interrupted the question before the answer categories were read aloud. For D1a and D1b, this occurred 20.0 percent and 6.9 percent of the time, respectively. For

item D1a, 10.5 percent of the time the question had to be repeated. For items D2a and D2b, which ask whether the incident involved unwanted kissing or groping, respondents answered inadequately 18.2 percent and 12.7 percent of the time by not specifying if the person threatened to, tried to, or actually did kiss or grope them. Instead, respondents were usually just responding “Yes.” This led to respondents needing to change their response 18.9 percent and 9.8 percent of the time to a more specific answer that fit a category. These issues were not identified by the behavior coding for the other items in the series.

The series of questions asking whether respondents felt an emotion as a result of the incident for less than one month or one month or longer (G21) also seemed to be cognitively demanding. Respondents answered inadequately as much as 15 percent of the time and respondents changed their answer as much as 13 percent of the time.

All other behavior coding of items in the DIF showed respondents had little difficulty (less than 6% required clarification, less than 7% required the question to be repeated, and less than 11% changed their answer).

Vignettes. As described in the first section of this chapter, the interviewers read the scenarios and the follow-up questions exactly as worded the majority of the time. There was a lower rate of respondents providing adequate answers to some of the follow-up questions. These questions are listed below.

Alcohol vignette items

1. Has a male used force or threats of force to make Sue have vaginal sex against her will? By vaginal sex, it means putting his penis in her vagina.
2. Has anyone made Sue have sex when she was unable to consent because she was too drunk, high, or passed out?
3. Do you think Tom was trying to get Sue drunk so he could sexually take advantage of her?
4. Did the alcohol make Sue unable to give consent?
5. On a scale from 1 to 10 where 1 means Sue definitely did not give consent and 10 means she definitely gave consent, to what extent did Sue give consent?

Coercion vignette items

1. Has a male used force or threats of force to make Becky have vaginal sex against her will?
2. Has anyone made Becky have any type of sex against her will by threatening to cause problems for her, such as at her job or school, at home, in her relationships, or in any other way?
3. On a scale from 1 to 10 where 1 means Becky was definitely not pressured and 10 means she was definitely pressured, to what extent was Becky pressured to have sexual intercourse?

For the alcohol questions 3-5 and question 3 for the coercion question, 20 to 35 percent provided an inadequate answer.

When respondents didn't provide an adequate answer, many provided qualified answers indicating doubt.

Results – Volunteer Sample

The volunteer sample offers an interesting contrast to the general population sample because, unlike the GP respondents, the VO respondents initiated their contact with the study. While they did not know about the survey content at the time, we can assume that their reactions to the survey could be different from those in the general population who were called by the study team and were asked to participate. The VO sample was also quite a bit younger than the GP sample, ranging in ages between 18 – 29 years old.

Consent Form. The VO behavior coding results on the consent form were similar to the GP results. The consent form was administered with little difficulty in the VO sample.

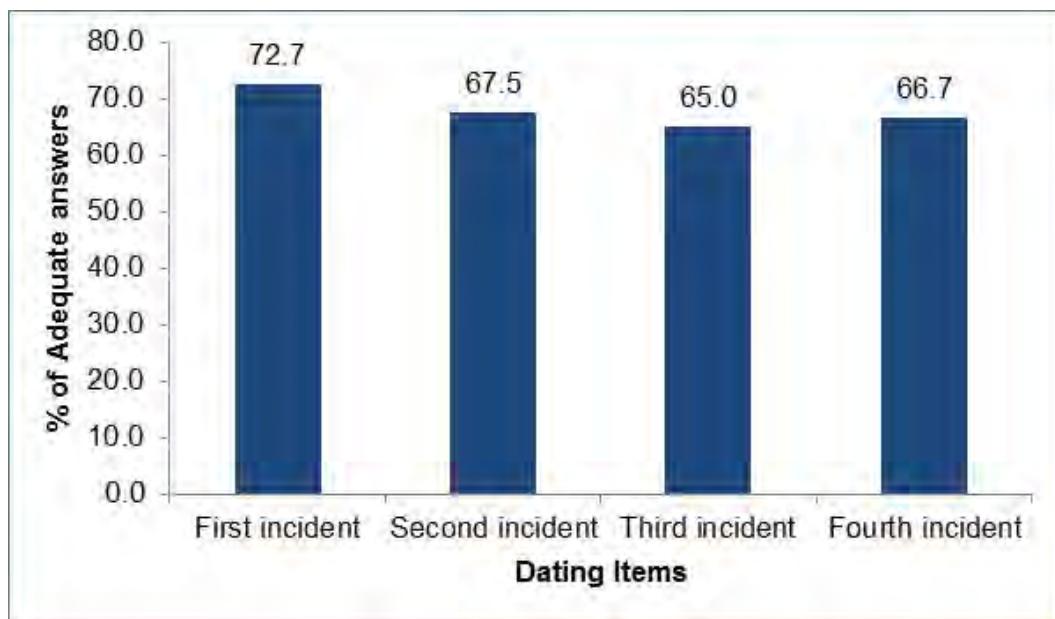
Demographics. Similar to the GP sample, a lower percentage of respondents' answers were adequate for the income question (78.7%) and the question on how long the respondent has been at their address (69.3%).

Sexual Victimization Screener. Among the VO sample, similar issues as described for the GP sample were found with the victimization screener and detailed incident form. Consistent with GP, the de-duplication of incidents worked well in the screener section. All respondents answered

adequately and respondents needed the question repeated or required clarification less than 4 percent of the time. Similar to GP respondents, VO respondents had greater difficulty providing the month and year of incidents prior to the most recent incident. Respondents adequately dated the first incident by providing the month and year (e.g., June 2014) 72.7 percent, the second incident 67.5 percent, the third incident 65 percent, and the fourth incident 66.7 percent of the time (figure 11-2). Respondents changed their answer to dating the first incident 11.4 percent of the time.

One exception was for the sexual victimization screener questions. In contrast to GP, VO respondents didn't interrupt or change answers to the first three sexual victimization screener items. This might be expected, given both that the VO sample was quite a bit younger than the GP sample, as well as the fact they were volunteers, with different motivations to take the survey.

Figure 11-2. Percent of adequate answers by victimization screener incident dating items for females age 18-29 in the volunteer sample (unweighted)



Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Behavior and Tactics. In comparison to the GP, more issues from the behavior coding were identified for the VO sample in the detailed incident form. These issues came about specifically with the items D2a, D2b, and D2c, which ask whether the incident involved unwanted kissing, groping, or exposing sexual body parts. Respondents answered questions regarding kissing, groping, and exposing sexual body parts inadequately 24.4 percent, 28.4 percent, and 12.3 percent of the time and changed their answer to these questions 19.5 percent, 22.2 percent, and 11.1 percent of the time. In most of these cases, inadequate answers consisted of providing an ambiguous response like

“Yes.” In these instances, it wasn’t clear if a respondent was saying the offender threatened to, tried to, or actually did do a certain act. Items D1a, D1b, D1c, and D1d, which ask about penetrative acts, had low rates of inadequate answers. Finally, when respondents were asked how long they felt sad, vulnerable, violated, and untrusting as a result of the incident (G21), they answered inadequately 13 to 38 percent of the time (i.e., saying “yes” or “moderately” instead of “for one month or longer”) and changed their answer as much as 29 percent of the time.

Vignettes. Although interviewers in the VO sample read the “stories” and vignette follow-up questions exactly as worded the majority of the time as in the GP sample, VO respondents answered the questions adequately more often. Across all five alcohol vignette follow-up items, 83 to 96 percent of respondents’ answers were adequate; 82 to 97 percent of the response were adequate across the three coercion vignette items.

11.3.2 CATI Interviewer Effects

One of the issues associated administering a sensitive survey like the RSA Pilot Test is the performance of the interviewers. While all attempts were made to minimize interviewer variance by using a standardized protocol and intensive interviewer training, interviewers vary in their ability to adhere to the protocols. In addition, interviewers bring with them their own style and personality when administering questions. Previous research has shown that interviewer effects tend to be largest for survey questions that are open-ended, difficult to answer, require follow-up probing, or which otherwise require the interviewer to exercise discretion in order to obtain an adequate answer (Fowler & Mangione, 1990). As noted elsewhere, the behaviorally specific questions are complex. When one considers the sensitivity of the questions as well, one might expect interviewers to vary in how they administer the questions. However, there is little evidence that suggests items with sensitive content such as sexual victimization are especially subject to interviewer effects (Fowler & Mangione, 1990), although there has not been a great deal of research on this topic.

To examine interviewer effects, it is necessary for the design to be interpenetrated with respect to assignments. That is, sample cases should be randomly assigned to interviewers. If this is not done, then it is difficult, if not impossible, to separate out the effects of the type of sample that is administered for the survey from the effects of the interviewer. For the in-person survey, assignments are made on the basis of geographic proximity to the interviewer’s residence to minimize travel costs. Since geography is linked to victimization rates, this makes it difficult to estimate interviewer effects for the ACASI survey. However, for the RDD portion of the CATI

survey, the cases are assigned in a quasi-random fashion. There is an automated algorithm that assigns each sampled telephone number to the interviewer who is available at the time. For example, Groves and Magilavy (1986) estimate interviewer effects for a series of RDD surveys assuming this assignment is random.

In order to examine the extent to which victimization prevalence rate varied within and between interviewers, intraclass correlation coefficients for the interviewer (ICCs) were calculated for several types of sexual victimization. A series of multivariate hierarchical regression models were estimated for the general population sample, predicting sexual victimization prevalence rates. The types of sexual victimization that were predicted in these models were rape, sexual assault, and other unwanted sexual contact. These models provide the proportion of variance that is accounted for between interviewers, which was used to calculate the ICC for interviewers.

The first level of the model predicted victimization for each respondent in the sample. The second, or interviewer, level of the model predicted the intercept term of the level 1. The intraclass correlation coefficient was then calculated by taking the estimate of the variance in the intercept (from the level 2 model) and dividing it by the total variance estimated from the model (Snijders & Bosker, 1999). The result is the proportion of variance in the victimization estimates that is accounted for at the interviewer level.

Out of 145 CATI interviewers who completed at least one interview with a general population respondent ages 18-49, 125 of them completed at least five interviews and were included in the analysis. From these 125 interviewers, 2,644 completed interviews were included in the models, and the median number of general population interviews completed per interviewer was 16.3.

Models were estimated in two different ways. The first set of models did not include any control variables at all (table 11-14). Results from these models showed that approximately 1 to 3 percent of the variance in victimization estimates is at the interviewer level. Intraclass correlation coefficients were significantly different from zero for rape (2.6%), sexual assault (1.8%), and other unwanted sexual contact (1.7%), when not controlling for any other variables. The magnitude of these are in the middle of the range reported in several of the studies cited above.

Table 11-14. Parameter estimates and intraclass correlation coefficients from multi-level regression models predicting prevalence of rape, sexual assault and other unwanted sexual contact with respondent characteristics and interviewer experience, for females ages 18-49 in the general population, 2014–2015

Parameter	Rape ^a		Sexual assault ^b		Other unwanted sexual contact ^c	
	Empty model	Full model	Empty model	Full model	Empty model	Full model
Intercept	0.033 †	0.054 †	0.058	0.076 †	0.068	0.103 †
Interview-level (level 1)						
Respondent age						
18-21*						
22-24	~	-0.002	~	0.001	~	-0.029
25-29	~	-0.024	~	-0.039	~	-0.063 †
30-34	~	-0.053 †	~	-0.065 †	~	-0.080 †
35-39	~	-0.045 †	~	-0.062 †	~	-0.105 †
40-49	~	-0.054 †	~	-0.070 †	~	-0.101 †
Respondent race/Hispanic origin	~		~		~	
White only ^{d,*}	~		~		~	
Black only ^d	~	-0.011	~	-0.023	~	-0.010
Hispanic	~	-0.023 †	~	-0.041 †	~	-0.054 †
Other ^{d,e}	~	-0.012	~	-0.014	~	-0.022
Respondent education						
Less than high school*						
High school/GED	~	0.002	~	0.016	~	0.007
Some college/associate's	~	0.003	~	0.023	~	0.039 †
Bachelor's or more	~	0.003	~	0.021	~	0.028
Respondent marital status						
Married*						
Not married	~	0.026 †	~	0.051 †	~	0.053 †

Table 11-14. Parameter estimates and intraclass correlation coefficients from multi-level regression models predicting prevalence of rape, sexual assault and other unwanted sexual contact with respondent characteristics and interviewer experience, for females ages 18-49 in the general population, 2014–2015 (continued)

Parameter	Rape ^a		Sexual assault ^b		Other unwanted sexual contact ^c	
	Empty model	Full model	Empty model	Full model	Empty model	Full model
Interviewer day in field						
Days 1-3*						
Days 1-3*						
Days 4-15	~	0.002	~	0.013	~	0.019
Days 16-25	~	-0.004	~	-0.003	~	0.030
Days 26+	~	0.004	~	0.016	~	0.016
Interviewer-level (level 2)						
Total completed interviews						
6-15 total completes*						
16-68 total completes	~	0.006	~	-0.010	~	-0.020
69+ total completes	~	0.007	~	-0.010	~	-0.011
Intraclass correlation coefficient (ICC)	2.6% †	2.3%	1.8% †	1.2%	1.7% †	0.5%

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

~ Not applicable. Not included in model.

† For parameters: estimate significantly different from comparison group at the 95% confidence level. For ICC: covariance parameter estimate for random intercept is significantly different from zero at the 95% confidence level.

^aIncludes penetrative sexual contact using force or while unable to consent.

^bIncludes non-penetrative sexual contact using force or while unable to consent.

^cIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported, or sexual contact in which the behavior or tactic was not specified.

^dExcludes persons of Hispanic or Latina origin.

^eOther race includes American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and persons identifying as “other” race (CATI only) or two or more races.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

This model assumes that completed surveys were randomly assigned to interviewers. However, this may not be the case—for example, because interviewers may tend to work different times of the day and days of the week. Respondents available to complete the survey during the day could be different than those who are only available to complete the survey during evenings and weekends.

The second model attempts to control for possible non-random allocation by introducing interview-level (level 1) control variables including respondent demographics (age, race/ethnicity, education, marital status) and interviewer experience, as measured by the number of days the interviewer had been in the field up to that point of the particular interview. This model also introduces the interviewer-level (level 2) control variable of total number of interviews completed by the interviewer. Specifically, it is used to predict the level 1 intercept. A positive coefficient would indicate that those who completed the most interviews tended to get higher victimization rates, while the opposite is the case for a negative coefficient.

Once all of these respondent and interviewer characteristic variables are controlled for, the proportion of interviewer variance decreases to 2.3 percent for rape, 1.2 percent for sexual assault, and 0.5 percent for other unwanted sexual contact. None of these are different from zero once taking into account sampling variation. We also note that none of the interviewer-related variables are statistically significant. Neither the number of days the interviewer had participated in the study up to the point of the interview or the total number of interviews completed were statistically significant. Overall, therefore, the impact of the interviewer on measuring victimization does not seem large or at least not much different than has been observed on other telephone surveys.

12. Data Reliability

The RSA Pilot Test included a reinterview for a small subset of respondents. These data were used to evaluate the items on both the SV screener and the detailed incident form. The first section of this chapter provides the design of the reinterview. The second section describes the results for the SV screener. The third section describes the reliability for selected items on the detailed incident form. The final section summarizes the results.

The RSA Pilot Test selected a subsample of respondents to participate in a reinterview several weeks after completing the interview. Reinterview studies for test-retest reliability are commonly used for assessing the quality of data from federal surveys (Brick, Rizzo, & Wernimont, 1997; Graham, 1977; Singer & Ennis, 2003; Substance Abuse and Mental Health Services Administration, 2010). Reinterview studies are intended to measure the reliability or consistency of survey responses over time. This provides a direct measure of the test-retest correlation.

The research questions addressed by this analysis are —

1. How reliable are the items on the sexual victimization screener?
2. How reliable are the measures of victimization from the detailed incident form (DIF)?
3. How does reliability vary by mode of interview?

12.1 Methodology

Reinterview studies use the correlation between a survey measure administered at time t and time $t + 1$ as an estimate of the test-retest reliability of that survey measure. This assumes the following model of the measurement process shown in (1) where the response obtained for the i th respondent (y_{it}) at time t equals a true value (X_i) plus a deviation (e_{it}) from the true value for the i th person at time t .

$$(1) \quad y_{it} = X_i + e_{it}$$

- y_{it} = response obtained for the i th person at time t
 X_i = true value of the measurement for the i th person
 e_{it} = response deviation from the true value for the i th person at time t

This model of the measurement process makes two key assumptions in order to estimate test-retest reliability with the correlation between two survey measures over time.

$$(2) E(e_{i1}) = E(e_{i2}) = 0$$

$$(3) \text{Cov}(e_{i1}, e_{i2}) = 0$$

First, the assumption (2) is that the expected value of the errors at both points in time is zero. This means the expected value of the survey measure at both points in time is X_i . The second assumption states that the errors between measurement occasions are not correlated. These assumptions can be violated if the respondent's true value changes between measurement occasions. For example, the respondent may be victimized between the two interviews. The respondent may also change her answer to the second interview because the first interview made her think differently about the answer to a question. These two circumstances would lead to lower estimates of reliability. Alternatively, respondents may also remember their answer to the first interview when answering questions during the second interview. This would lead to a higher estimate of reliability than if memory effects were not present.

12.1.1 Sampling and Response Rates

The reinterview sample was restricted to women ages 18-49. Respondents were selected for the reinterview based on their responses to the SV screener. Because the sample size for the reinterview was relatively small and the incidence of reporting a victimization is low (e.g., 5% of victims report a rape or sexual assault), the sampling for the reinterview concentrated on those reporting a victimization. In order to be selected for the reinterview, respondents had to report an incident in either the past 12 months or within their lifetime. We selected respondents from three groups. The first group consisted of respondents who reported a past 12-month incident to one of the first eight items on the sexual violence screener (SV1-8). These items include measures of unwanted penetration. The second group consisted of respondents who reported a past-12 month incident to items 9-14 of the sexual violence screener (SV9-14). These items cover unwanted sexual touching. The third group reported at least one lifetime incident, but no incidents in the past 12 months.

The reinterview selection rates started out being higher for the respondents who reported incidents over the past 12 months. Selection rates varied over the course of the field period

depending on progress toward the target number of reinterviews. Generally, the selection rates for respondents with only lifetime reports were increased over the course of the field period to meet the targeted number of reinterviews.

Appendix I provides detailed selection and response rates for the four different groups (2 modes x 2 sample types = 4). The reinterview response rates were 63.3 percent and 60.2 percent for the ACASI and CATI general population samples, respectively. For the volunteer sample, they were 86 percent for the ACASI and 50.4 percent for the CATI. These rates are a percent of those who were selected for the reinterview, so they include both those who did not agree to do the interview as well as those who did not complete among those who initially agreed. There was some tendency of the CATI respondents who reported a victimization at the first interview to both not agree to the reinterview and to not complete the reinterview if they had initially agreed. This may be due to the greater amount of time it took to fill out the detailed incident form (see Chapter 6).

The composition of the three sampling strata (SV1-8, SV9-14, and lifetime) differed by the four groups. Approximately 80 percent of the CATI general population respondents reported only lifetime victimization in the first interview, whereas only one half of the ACASI general population reinterview respondents reported only lifetime victimization. In the volunteer samples, ACASI reinterview respondents reported a higher percentage of victimizations among the first eight screener items during the main interview compared to the CATI volunteer reinterview respondents.

To minimize these different sources of error, reinterviews were targeted for approximately 2 weeks after the main interview. The survey was reasonably successful at this. Approximately 80 percent of the reinterviews were completed within 3 weeks of the main interview (Appendix I).

12.1.2 Primary Measure of Reliability

Many of the analyses discussed below involve agreement between binary variables indicating whether or not a respondent experienced a particular type of victimization. It would be possible to report the simple agreement rate between the two interviews; however, simple agreement rates tend to be very high for rare events like the reports of rape and sexual assault. There is a very high probability that the reports between interviews will agree by chance. Kappa is measure of reliability

for categorical variables that accounts for chance-agreement. “Chance-agreement” is computed by assuming the results of the two interviews are independent. Kappa is computed using equation 1.

$$(1) \quad K_c = \frac{p_a - p_e}{1 - p_e}$$

Where p_a is the proportion of the code assignments where the two survey reports agree and p_e is the chance-agreement probability for the two survey reports. The numerator in the formula above represents the percent of cases that agree beyond chance. The denominator represents the percent of cases for which one would not expect any agreement by chance. Landis and Koch (1977) provide a rubric for interpreting kappa:

- poor agreement (kappa less than 0)
- slight agreement (kappa between .00 and .20)
- fair agreement (kappa between .21 and .40)
- moderate agreement (kappa between .41 and .60)
- substantial agreement (kappa between .61 and .80)
- almost perfect agreement (kappa between .81 and 1.0).

12.1.3 Reinterview Weights and Statistical Testing

A reinterview weight was constructed to account for the differential rates at which respondents from the three victimization groups (SV1-8, SV9-14, and lifetime) were selected, as well as nonresponse to the reinterview within each strata. Separately for the GP and VO samples, the ratio of the number of original interviews to the number of completed reinterviews was calculated for each victimization group. For the VO sample, the reinterview weight was simply this ratio. For the GP sample, the product of the ratio and the equalized final weight was calculated; replicate weights were also created by multiplying the ratio by the replicate equalized weights. The reinterview weight was used in the analysis of the reliability of the screener questions.

For the analysis of the volunteer screener cases this weight was normalized by dividing each case weight by the mean of the weights. This weight was then used when conducting the statistical analyses for the screener, assuming a simple random sample. For the analysis of the general population screener cases the final weight and replicate weights for the first interview were

multiplied by the reinterview weight. The analysis of the general population cases was conducted using the *survey* package in the statistical software R to compute the correct weighted estimates and standard errors for the reliability statistics.

The analysis of the DIF items (section 3) combined the results from the GP and VO samples using unweighted data. Statistical tests were conducted assuming a simple random sample.

12.2 Results for the Screener

Table 12-1 summarizes the measures of the reliability for the screener items for the general population sample. The table summarizes reliability across groups of screener questions. More detailed results showing reliability estimates of the individual screener items can be found in Appendix I. The top panel of the table combines the general population sample from both modes of data collection. The bottom two panels of the table show reliability estimates for the ACASI and CATI modes of data collection separately. The table summarizes the reports to the two interviews and also displays the estimated kappa values for each group of screener questions.

The screener items are grouped by the types of incidents they were intended to measure: SV1-5 are the items for completed rape, SV6-8 for completed other unwanted penetration and attempted unwanted penetration, and SV9-14 for unwanted sexual touching and sexual non-contact.

Overall, the top panel of table 12-1 kappa coefficients for each group of screener items between .60 and .67. The kappa coefficient increases to .76 when considering any report within the past 12 months. The interior middle rows provide an idea of the direction of any inconsistencies. For items covering penetration (SV1-5; SV6-8), there are approximately the same percentage of respondents who initially reported in one of these items, but did not at the second interview (e.g., SV1-5 – 1.9%) and vice versa (e.g., SV1-5 – 2.9%). For sexual contact and non-contact, there is a tendency for more respondents to report something at the second interview and not at the first.

For the items covering completed rape, the ACASI has higher reliability than the CATI, although this difference is not statistically significant at the 5 percent level (0.71 vs. 0.43, $p < .10$). The opposite is the case for other unwanted penetration and attempted unwanted penetration (SV6-8).

Table 12-1. Estimates of reliability for the sexual victimization screener items overall and by mode in the general population, 2014–2015

Item	Time 1=Yes ^a		Time 1=No ^a		Kappa ^b
	Time 2=Yes ^a	Time 2=No ^a	Time 2=Yes ^a	Time 2=No ^a	
Overall					
Any past 12 month incident ^c	27.0 %	7.1 %	3.6 %	62.2 %	0.76
SV1 – SV5: Rape ^d	4.0	1.9	2.9	91.3	0.60
SV6 – SV8: Other unwanted sex ^e	7.2	3.4	3.3	86.0	0.64
SV9 – SV14: Sexual contact ^f	20.5	8.3	4.8	66.4	0.67
Lifetime incidents only ^g	55.6	10.3	3.8	30.8	0.70
ACASI^h					
Any past 12 month incident ^c	25.7 %	5.9 %	4.0 %	64.4 %	0.77
SV1 – SV5: Rape ^d	4.2	1.6	1.6	92.6	0.71
SV6 – SV8: Other unwanted sex ^e	5.8	3.9	3.4	86.9	0.57
SV9 – SV14: Sexual contact ^f	19.4	7.4	4.8	68.4	0.68
Lifetime incidents only ^g	56.4	12.0	2.5	29.0	0.69
CATIⁱ					
Any past 12 month incident ^c	29.8 %	9.6 %	2.7 %	57.9 %	0.73
SV1 – SV5: Rape ^d	3.6	2.5	5.5	88.4	0.43
SV6 – SV8: Other unwanted sex ^e	10.2	2.5	3.1	84.2	0.75
SV9 – SV14: Sexual contact ^f	22.6	10.1	4.9	62.4	0.64
Lifetime incidents only ^g	53.9	6.7	6.4	33.1	0.72

Note: Estimates are based on unweighted data. Standard errors for kappa statistics can be found in Appendix A.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview. “Yes” indicates the respondent reported experiencing that type of incidents in that interview. “No” indicates the respondent did not report experiencing that type of incident in that interview.

^bKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^cIncludes incidents that occurred within the past 12 months of the interview.

^dIncludes sexual victimization screening items about vaginal, oral, anal, and digital forced penetration and penetration while unable to consent due to alcohol or drugs.

^eIncludes sexual victimization screening items about coerced penetration, other unwanted penetration, and attempted penetration.

^fIncludes sexual victimization screening items about unwanted kissing, groping, attempted kissing or groping, sexual exposure, and participation in sexual photos or movies.

^gIncludes incidents that occurred at any point in the respondent’s lifetime.

^hAudio computer-assisted self-interview.

ⁱComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The reporting of any 12-month incident has similar reliabilities for both modes, as does reporting a lifetime incident. For ACASI, the inconsistencies are different for lifetime victimization. Those reporting only a lifetime victimization at the first interview were less likely to be in this category at the second interview (12.0% vs. 2.5%). The respondent may have reported a 12-month event at the second interview instead or no victimization at all.

A similar level of reliability was found for the volunteer sample (Appendix I).

The individual items (Appendix I) generally have lower reliabilities. Most are between 40 and 60 percent. The ACASI has higher reliabilities for the forced penetration items, although the sample sizes are too small to reliably estimate these same items for the CATI. For both modes, the item on alcohol and drug facilitated incidents has a reliability of around 71 percent. Attempted penetration (SV8) has a reliability of 57 percent.

To get a sense of how respondents may have changed between interviews, responses were coded by prioritizing the screener items that were endorsed. Setting a priority was necessary because more than one screener item could be endorsed at each interview. The section on the DIF below examines shifts at an incident level. The priority used for this coding, from highest to lowest, was rape (SV1-5), other unwanted penetration and attempted penetration (SV6-8), unwanted sexual contact (SV9-14), and lifetime only incidents. For the second interview, “no victimizations” was also included. Among the respondents who endorsed the completed rape items at the first interview, 68.3 percent did so at the second interview (table 12-2). The remaining 32 percent were evenly divided among the other categories at the second interview. The patterns are different for the other types of screener items. For other items on unwanted penetration and attempted penetration, 45.9 percent reported this at both interviews, while 21.9 percent reported completed rape instead. About as many endorsed the items on unwanted sexual contact and non-contact (18.3%). For unwanted sexual contact and non-contact, 62.3 percent endorsed these items at both interviews, while about 25 percent either just endorsed the items on a lifetime incident or no incidents at all at the second interview. These patterns are very similar by mode, as well as the VO sample (Appendix I).

Table 12-2. Type of sexual victimization screener incident reported in second interview, conditional on type of victimization reported in first interview, for the general population

Time 2 ^a	Time 1 ^a			
	Past 12 month incidents ^b			
	SV1 – SV5: Rape ^c	SV6 – SV8: Other unwanted sex ^d	SV9 – SV14: Sexual contact ^e	Lifetime incidents only ^f
Past 12 month incidents ^b				
SV1 – SV5: Rape ^c	68.3 %	21.9 %	2.7 %	0.7 %
SV6 – SV8: Attempted Penetration and other unwanted sex ^d	9.6	45.9	8.8	1.2
SV9 – SV14: Sexual contact ^e	9.6	18.3	62.3	3.5
Lifetime incidents only ^f	5.3 %	12.1 %	12.4 %	84.4 %
No incidents reported	7.3 %	1.8 %	13.9 %	10.1 %
Number of unweighted sample cases	48	74	167	574

Note: Estimates are based on unweighted data.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview. "Yes" indicates the respondent reported experiencing that type of incidents in that interview. "No" indicates the respondent did not report experiencing that type of incident in that interview.

^bIncludes incidents that occurred within the past 12 months of the interview.

^cIncludes sexual victimization screening items about vaginal, oral, anal, and digital forced penetration and penetration while unable to consent due to alcohol or drugs.

^dIncludes sexual victimization screening items about coerced penetration, other unwanted penetration, and attempted penetration.

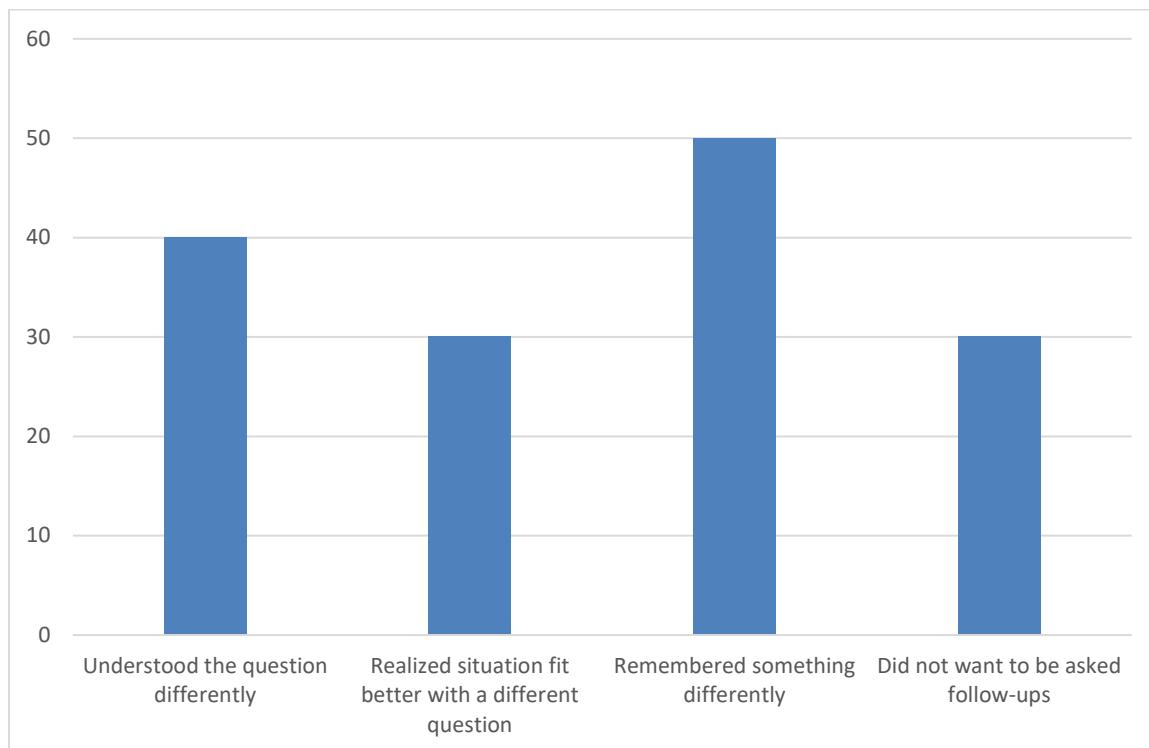
^eIncludes sexual victimization screening items about unwanted kissing, groping, attempted kissing or groping, sexual exposure, and participation in sexual photos or movies.

^fIncludes incidents that occurred at any point in the respondent's lifetime.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

To get a sense of why respondents may have changed answers, the re-interview asked respondents if they remembered changing their answers to particular sections of the questionnaire. ACASI respondents were asked whether they changed their answers to the 12-month victimization screener items (figure 12-1). Approximately 25 percent of the respondents remember changing at least one of their answers to the victimization screener items ($n=123$). Approximately 40 percent of these respondents said that they “understood the questions differently” at the second interview. Similarly approximately 30 percent said they “realized my situation fit better with a different question on the survey.” These are both indications of respondent conditioning—that is, respondents learned something from the first interview and made adjustments to their answers at the second interview. Combining these two answers together, approximately 60 percent of the respondents were in one of these two categories. A second common response was that the respondent “remembered something differently” (50% of those that changed). About 30 percent of respondents declined to say they had been a victim because they “did not want to be asked the detailed follow-up questions.”

Figure 12-1. Reasons for changing answers to 12-month victimization screener items (ACASI)*

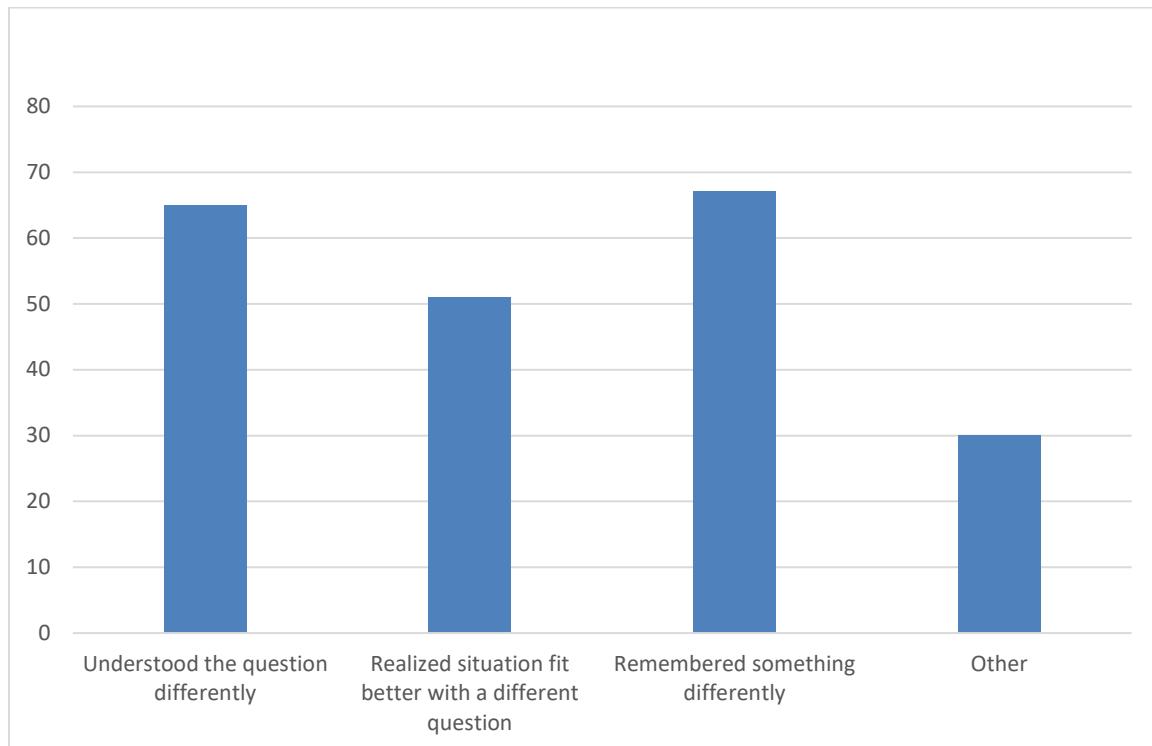


* Percent of respondents who said they changed an answer to the 12-month victimization questions ($n=123$).
Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

A similar pattern was found on the CATI, which asked about lifetime victimization first and then 12 month victimization.⁴⁷ About 25 percent of these respondents said they changed their answers to at least one of the lifetime screener questions (n=166). A high percentage of respondents said they “understood the questions differently” (61%) or their “situation fit better with a different question on the survey” (51%). A significant number also said that they “remembered something differently” (67%). Almost a third of the respondents selected the “other specify” response option and expressed a mix of all three reasons noted above. One sentiment that was new in these open-ended responses was an increase in trust in the interviewer in the second interview.

“Because I understand the questions better now and trust the person doing the survey...and feel more comfortable answering the questions...”

Figure 12-2. Reasons for changing answers to lifetime victimization screener items (CATI)*



* Percent of respondents who said they changed an answer to the lifetime victimization questions (n=166).
Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

⁴⁷The CATI version of the survey first asked whether there was a lifetime victimization and then followed up each ‘yes’ response with whether anything happened in the last 12 months.

These data should be interpreted with some caution because they rely on a self-report of whether the respondent changed an answer. We can see from the data above, for example, that while 25 percent of the respondents said they changed at least one answer to a screener question, the proportion that actually changed was more than this. Nonetheless, these data do provide a window into some of the reasons why changes occurred. The predominant reason for change in the screener items was respondents reconsidering their original responses, based on what they learned in the first interview, remembered something different between interviews, or were just more comfortable with reporting something the second time around. Regardless of which of these three occurred, it points to respondent difficulties with answering the questions in the first interview. These difficulties are not necessarily harmful to the final estimates, as long as the incident is captured and a detailed incident form is filled out. But some of these errors could mistakenly place the incident out of (or in) scope for the survey. A secondary reason for change was to avoid answering the follow-up questions on the detailed incident form. On the one hand, this type of conditioning is not harmful for a one-time survey. Respondents won't have the opportunity to apply their knowledge of the consequences of their responses. But if the survey is done as part of a rotating panel design, like the current NCVS, it could contribute to underreports of events.

12.3 Reliability of the Detailed Incident Form

While the above analysis examine whether the patterns are consistent between interviews, it is not entirely clear whether respondents are changing the particular incidents that are reported or if they are answering the questions differently for the same incident. In this section, the DIF is used to first assess the proportion of incidents that were reported at both the first and second interviews. The second analysis examines the reliability of the key items on the DIF.

12.3.1 How Often is the Same Incident Reported at Both Interviews?

To assess how often respondents reported the same incident at both interviews, the analysis used both the narratives and the DIF to match incidents across interviews.

Matching DIFs with Narratives

Some respondents reported multiple incidents and filled out more than one incident form. To identify whether a DIF was referring to the same incident, the narratives were reviewed to match incidents. Two coders and one reviewer read all narratives provided by each respondent and compared each narrative from the first interview to the narratives provided in the second interview. There were a maximum of three narratives from the first interview that could be matched to three narratives from the second interview.

In order for narratives to be considered a match, at least one of the following criteria needed to be met:

- Both narratives included unique identifiers of the location or circumstances of the incident matched.
- Both narratives included a description of the behaviors in the incident that closely matched.
- Both narratives included a description of the offender or demographic characteristics that closely matched.
- Both narratives included a description of the tactic used in the incident that closely matched.

In some cases there were multiple narratives reported within an interview that were very similar to each other. If there was not any distinguishing information that allowed the similar incidents to be matched to a narrative in the second interview, no match was made. In other cases, the narrative in one of the interviews may have contained detailed information about the location, behaviors, tactics, or offender characteristics, but if the narrative from the other interview did not contain sufficient detail that allowed a match to confidently be made, then no match was made. Appendix I provides a few examples of how incidents were matched.

This matching was done for all respondents who had at least one narrative for both the first and second interview. This left incidents that could not be considered in this matching process. In order to decide if there were other incidents that matched across interviews, incidents without narratives were matched based on a statistical matching algorithm. The hand-coded matches with narratives were used to develop a statistical model predicting a match using the variables on the screener and the DIF. This model was then applied to the reports for which no narrative was given

to match those without a matched incident. Appendix I provides more details on the procedures to build and evaluate this model.

Because of the relatively small sample sizes, the analysis below uses unweighted data and combines across both mode of interview and type of sample (general population and volunteer). As noted in the screener section, there were very few differences, if any, between mode and sample type. If anything, the reliabilities for the volunteer sample were a bit lower than the general population, although these differences were not large.

How Many Incidents Match Between Interviews?

This analysis generated two estimates of whether an incident was reported at both interviews. The first is based on the percentage that were matched using just the narratives. The denominator for this match rate is all individuals who had at least one narrative for both interviews. These are the incidents that could have been matched using this method. The overall match rate for this is 66 percent. This estimate is likely too low because it excludes those incidents for which there is no narrative. If these are included using the statistically matched incidents, the overall match rate goes up to 72 percent. This estimate may be too high because the number of false positives created when statistically matching incidents without a narrative (see Appendix I).

There were more incidents reported at the reinterview than at the first interview. This may partly reflect the additional time period associated with the second interview. Respondents may have experienced an incident between the initial interview and the reinterview, which would be reported at the second interview.

12.3.2 How Reliable are Measures of Behaviors and Tactics on the DIF?

This section provides data on the reliability of selected questions on the DIF. Using the incidents that were matched using the narratives, the responses to key items on the DIF were compared between the first and second interviews. This section discusses the items asking about unwanted behaviors, the use of force, and inability to consent – three key sets of items related to classifying the incident.

Unwanted Behavior Items

There were two sets of unwanted behavior items. Items D1a through D1d asked about different forms of unwanted penetration, including vaginal, oral, anal, and digital. The second set (D2a – D2d) asked about unwanted touching, groping, and sexual non-contact.

For unwanted penetration, tables 12-3 to 12-6 compare the responses from each interview for the incidents that were matched based on the narratives. Kappa coefficients and the percent that agree are also shown in each table. The sample sizes for any particular type of penetration is low, but the same pattern is evident for each. The Kappa coefficients are in the moderate range (55% – 60%). There is relatively high agreement when respondents say that no unwanted penetration occurred. Approximately 90 percent of those reporting that “vaginal penetration did not occur” agree across interviews. There is somewhat less agreement for completed penetration. For example, for vaginal penetration, out of the 32 respondents who said that vaginal penetration occurred at the initial interview, 24 (75%) said it happened at the reinterview. Almost all of the discrepancy is switching to the “Did not happen” category. For vaginal penetration, the change is symmetrical between interviews—approximately the same number switch from completed to did not happen from time 1 to time 2 as switch in the other direction from time 2 to time 1 (e.g., 7 from time 1 to time 2; 8 from time 2 to time 1). This results in about the same number of completed vaginal incidents reported at each time. The change is not quite as symmetrical for the other types of penetration, with slightly more incidents being reported as completed at time 1 than time 2, although this difference is not large.

The largest discrepancies are for the threatened and attempted categories. The percentage agreeing between the interviews is around 30 percent for threatened and slightly higher for attempts. As with the completed penetrations, most of those that don’t agree selected the “Did not happen” category. As with vaginal penetration, the change is symmetrical. For the other types of penetration, there are slightly more completed acts being reported at time 1.

Table 12-3. Agreement rates and kappa statistic for detailed incident form question on unwanted vaginal penetration (D1a), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a				Did not happen	Time 1 ^a agreement rate ^e
		Threatened ^b	Attempted ^c	Completed ^d	Time 2 ^a		
Overall kappa ^f	0.56						
All incidents	338	18	28	33	259		
Threatened ^b	23	8	3	1	11	34.8 %	
Attempted ^c	25	3	14	0	8	56.0 %	
Completed ^d	32	0	1	24	7	75.0 %	
Did not happen	258	7	10	8	233	90.3 %	
Time 2 agreement rate ^g		44.4 %	50 %	72.7 %	90 %	82.5 %	

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^cIncludes incidents where the offender physically tried, but did not complete the behavior.

^dIncludes incidents where the offender completed the behavior.

^eIndicates percent of Time 1 reports that are the same as Time 2 reports.

^fKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^gIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

Table 12-4. Agreement rates and kappa statistic for detailed incident form question on unwanted oral penetration (D1b), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a				Time 1 ^a agreement rate ^e
		Threatened ^b	Attempted ^c	Completed ^d	Did not happen	
Overall kappa ^f	0.55					
All incidents	338	6	6	17	309	
Threatened ^b	11	3	1	0	7	27.3 %
Attempted ^c	11	1	3	3	4	27.3 %
Completed ^d	20	0	2	11	7	55.0 %
Did not happen	296	2	0	3	291	98.3 %
Time 2 agreement rate ^g		50.0 %	50.0 %	64.7 %	94.2 %	91.1 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^cIncludes incidents where the offender physically tried, but did not complete the behavior.

^dIncludes incidents where the offender completed the behavior.

^eIndicates percent of Time 1 reports that are the same as Time 2 reports.

^fKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^gIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 12-5. Agreement rates and kappa statistic for detailed incident form question on unwanted anal penetration (D1c), 2014-2015

Time 1 ^a	All incidents	Time 2 ^a				Time 1 ^a agreement rate ^e
		Threatened ^b	Attempted ^c	Completed ^d	Did not happen	
Overall kappa ^f	0.53					
All incidents	340	2	2	6	330	
Threatened ^b	3	1	0	0	2	33.3 %
Attempted ^c	4	1	0	0	3	0.0 %
Completed ^d	9	0	1	5	3	55.6 %
Did not happen	324	0	1	1	322	99.4 %
Time 2 agreement rate ^g		50.0 %	0.0 %	83.3 %	97.6 %	96.5 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^cIncludes incidents where the offender physically tried, but did not complete the behavior.

^dIncludes incidents where the offender completed the behavior.

^eIndicates percent of Time 1 reports that are the same as Time 2 reports.

^fKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^gIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 12-6. Agreement rates and kappa statistic for detailed incident form question on unwanted digital penetration (D1d), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a				Time 1 ^a agreement rate ^e
		Threatened ^b	Attempted ^c	Completed ^d	Did not happen	
Overall kappa ^f	0.58					
All incidents	338	4	10	20	304	
Threatened ^b	5	0	1	0	4	0.0 %
Attempted ^c	17	0	6	2	9	35.3 %
Completed ^d	24	0	2	16	6	66.7 %
Did not happen	292	4	1	2	285	97.6 %
Time 2 agreement rate ^g		0.0 %	60.0 %	80 %	93.8 %	90.8 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^cIncludes incidents where the offender physically tried, but did not complete the behavior.

^dIncludes incidents where the offender completed the behavior.

^eIndicates percent of Time 1 reports that are the same as Time 2 reports.

^fKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^gIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

For unwanted touching and non-contact, the agreement rates are slightly lower. Tables 12-7 and 12-8 provide the data for the two most prevalent unwanted behaviors: kissing and groping. The number of incidents involving the non-contact questions was too small to break out on an individual basis. The overall agreement rates are around 70 percent. The Kappa statistics are .54 for kissing and .49 for groping. As with unwanted penetration, the attempts and threats have the lowest agreement rates. However, unlike penetration, many of the changes for the threats and attempts move to the completed category at the other interview. For example, for unwanted kissing, of the 66 respondents who reported an attempt at the first interview, 14 said it was completed at the second interview, and 17 said it did not happen at the second interview. For groping, more of those reporting an attempt at the first interview reported it as being completed (24) compared to it not happening (14).

Overall, the inconsistencies identified above point to problems with the unwanted behavior items, especially for incidents classified as an attempt or threat. This is consistent with the inconsistencies identified in Chapter 7, which compared the screener responses to those on the DIF.

Table 12-7. Agreement rates and kappa statistic for detailed incident form question on unwanted kissing (D2a), 2014-2015

Time 1 ^a	All incidents	Time 2 ^a				Time 1 ^a agreement rate ^e
		Threatened ^b	Attempted ^c	Completed ^d	Did not happen	
Overall kappa ^f	0.54					
All incidents	339	11	55	100	173	
Threatened ^b	10	4	3	2	1	40.0 %
Attempted ^c	66	3	32	14	17	48.5 %
Completed ^d	103	2	11	69	21	67.0 %
Did not happen	160	2	9	15	134	83.8 %
Time 2 agreement rate ^g		36.4 %	58.2 %	69.0 %	77.5 %	70.5 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^cIncludes incidents where the offender physically tried, but did not complete the behavior.

^dIncludes incidents where the offender completed the behavior.

^eIndicates percent of Time 1 reports that are the same as Time 2 reports.

^fKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^gIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 12-8. Agreement rates and kappa statistic for detailed incident form question on unwanted groping (D2b), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a				Time 1 ^a agreement rate ^e
		Threatened ^b	Attempted ^c	Completed ^d	Did not happen	
Overall kappa ^f	0.49					
All incidents	337	4	52	165	116	
Threatened ^b	6	0	3	2	1	0.0 %
Attempted ^c	63	1	24	24	14	38.1 %
Completed ^d	163	1	17	125	20	76.7 %
Did not happen	105	2	8	14	81	77.1 %
Time 2 agreement rate ^g		0.0 %	46.2 %	75.8 %	69.8 %	68.2 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^cIncludes incidents where the offender physically tried, but did not complete the behavior.

^dIncludes incidents where the offender completed the behavior.

^eIndicates percent of Time 1 reports that are the same as Time 2 reports.

^fKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^gIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Use of Force

There were four items asking about use of force (D4a – D4d). Each question asks whether the individual did any of the following during the incident:

- hold you or pin you so you had difficulty moving
- use a weapon or threaten to use a weapon
- physically attack you or threaten to attack you, but not with a weapon
- physically attack or threaten to attack someone else
- use any other type of force.

Generally there was higher agreement for these items than for the type of unwanted behavior (i.e., questions D1 and D2). The reliabilities ranged from 70 to 80 percent. However, the prevalence of any of the behaviors reported for items b through d above is very low. The primary item that was endorsed was for holding or pinning down, which had an agreement rate of 88 percent overall and a reliability of .74 (table 12-9). Of those who endorsed this item at the first interview, 82 percent also endorsed at the reinterview. About as many people changed their endorsement of this item at each interview.

Table 12-9. Agreement rates and kappa statistic for detailed incident form question on holding or pinning down (D4a), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a		Time 1 ^a agreement rate ^b
		Yes	No	
Overall kappa ^c	0.74			
All incidents	305	105	200	
Yes	107	88	19	82.2 %
No	198	17	181	91.4 %
Time 2 ^a agreement rate ^d		83.8 %	90.5 %	88.2 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIndicates percent of Time 1 reports that are the same as Time 2 reports.

^cKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^dIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

Ability to Consent and Signs of Intoxication

Several items were used to classify an incident involving inability to consent. One was whether the respondent was passed out for all or part of the incident. This had four response categories—passed out for all of the incident, for part of the incident, not passed out at all, or not sure. This item has high agreement and reliability. Very few respondents reported being either totally passed out or not sure (table 12-10). The agreement rate for the response most commonly endorsed (not passed out at all) is 96 percent. For the other category that was endorsed with some frequency (passed out part of the time), the rate was 89 percent for the interview 1 match rate. The main inconsistencies for the latter were related to deciding on whether it was some of the time or whether the respondent was sure or not.

Table 12-10. Agreement rates and kappa statistic for detailed incident form question on whether passed out (G10), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a				Time 1 ^a agreement rate ^b
		Entire time	Part of the time	Not passed out	Not sure	
Overall kappa ^c	0.77					
All incidents	122	1	23	95	3	
Entire time	1	0	1	0	0	0.0 %
Part of the time	19	0	17	2	0	89.5 %
Not passed out	97	1	2	93	1	95.9 %
Not sure	5	0	3	0	2	40.0 %
Time 2 ^a agreement rate ^d		0.0 %	73.9 %	97.9 %	66.7 %	91.8 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIndicates percent of Time 1 reports that are the same as Time 2 reports.

^cKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^dIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

A second item related to the classification was whether the respondent could remember what happened (table 12-11). This also had four categories, with a scale of all, most, a few parts, or none of the incident. If the respondents said she did not remember any of the incidents, she skipped out of the rest of the alcohol questions. There were very few respondents who could not remember anything. The percent agreement was 80 percent and reliability of .65, with most of the movement between adjacent categories. For example, of those who originally said “all of the incident” and changed answers at the reinterview, most responded with the next point on the scale of “most of the incident.”

Table 12-11. Agreement rates and kappa statistic for detailed incident form question on whether respondent can remember what happened (G11), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a				Time 1 ^a agreement rate ^b
		All	Most	A few parts	None	
Overall kappa ^c	0.65					
All incidents	120	66	33	19	2	
All	73	61	9	3	0	83.6 %
Most	31	5	22	4	0	71.0 %
A few parts	13	0	2	11	0	84.6 %
Nothing	3	0	0	1	2	66.7 %
Time 2 ^a agreement rate ^d		92.4 %	66.7 %	57.9 %	100.0 %	80.0 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIndicates percent of Time 1 reports that are the same as Time 2 reports.

^cKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^dIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The other main criteria for classification as a victim of RSA was the victim was unable to provide consent because of alcohol or drugs. Respondents were very consistent on this question between the original and reinterview (table 12-12). Approximately 95 percent of the answers agreed, and this had a reliability of .87.

Table 12-12. Agreement rates and kappa statistic for detailed incident form question on whether alcohol or drugs made respondent unable to give consent (G12a), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a		Time 1 ^a agreement rate ^b
		Yes	No	
Overall kappa ^c	0.87			
All incidents	117	34	83	
Yes	32	30	2	93.8 %
No	85	4	81	95.3 %
Time 2 ^a agreement rate ^d		88.2 %	97.6 %	94.9 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIndicates percent of Time 1 reports that are the same as Time 2 reports.

^cKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^dIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Two items asked respondents about signs of intoxication (data not shown). One had to do with the ability to walk (D13) and another to communicate (G15). These two questions had rates of agreement of 85 to 90 percent and exhibited similar patterns. The percentage who reported showing signs of intoxication went down from the first interview to the reinterview. For example, 18 percent said they could not walk at the first interview and 13 percent said this at the reinterview. A similar pattern was observed for being able to communicate.

For incidents involving force or inability to consent to non-penetrative sexual contact, if the respondent said the perpetrator stopped immediately after she said “no,” the incident was classified as unwanted touching rather than sexual assault. There is 82.1 overall agreement for this item, with a reliability of 0.63 (table 12-13). The disagreement involves more respondents saying at the second interview that the perpetrator stopped immediately.

Table 12-13. Agreement rates and kappa statistic for detailed incident form question whether the perpetrator immediately stopped after saying “no” (G17), 2014–2015

Time 1 ^a	All incidents	Time 2 ^a		Time 1 ^a agreement rate ^b
		Yes	No	
Overall kappa ^c	0.63			
All incidents	196	87	109	
Yes	76	64	12	84.2 %
No	120	23	97	80.8 %
Time 2 ^a agreement rate ^d		73.6 %	89.0 %	82.1 %

Note: Estimates are based on unweighted data. Don't know and refuse responses at either Time 1 or Time 2 were not included in the analysis.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIndicates percent of Time 1 reports that are the same as Time 2 reports.

^cKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^dIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014–2015.

Overall Classification Using the DIF Items

This section provides the reliability of the classification algorithm. The algorithm combines the above items (e.g., unwanted behaviors; use of force; reports of ability to consent) to classify the victim into a particular type of unwanted sexual contact (see chapter 7 for a detailed description). The overall reliability of this measure is computed at both the person level and incident level. This section examines the reliability of everyone who was re-interviewed, not just incidents that were matched using the narratives.

Reliability of the Algorithm at the Person Level. Table 12-14 provides reliabilities and agreement rates for the general population sample for whether or not a respondent was classified into a particular type of at both interviews. These reliabilities are somewhat lower than for the screener. The overall category of rape and sexual assault had moderate reliability (.58). Separately, rape and sexual assault have reliabilities of .56 and .48, respectively. Overall, there is a general decline of “yes” responses at the second interview. For example, for rape and sexual assault victims at the first interview, 6.7 percent switched to a non-victim at the second interview, while 4 percent switched in the opposite direction. For the VO sample the inconsistencies in this direction are somewhat larger.

Nominally, the ACASI data have higher reliabilities for measurement of rape, but the difference is not statistically significant.

To examine how the types of incidents change for individual victims, a type of victimization measure was created that used the most serious victimization reported at each interview. The order of seriousness was rape (most serious), sexual assault, other unwanted contact, no unwanted behavior, lifetime incidents, and no incident reported at all (least serious). Table 12-15 cross-classifies this variable for the first and second interview for the general population sample.

One pattern of inconsistency is for victims to shift from a victimization to the lifetime-only category.⁴⁸ For example, of the 72 individuals reporting a rape at the first interview, 11 shifted to the “lifetime – not enough information” category. For this cell, there were no incidents in which respondents did not report enough information to classify the incident. This change is symmetrical with approximately the same number of respondents shifting from rape to lifetime as shifting from lifetime to rape between the interviews. These reflect someone reporting “yes” to a 12-month screener question and the incident being classified as a completed rape at one interview and just reporting a lifetime victimization at the other interview. Switching to a lifetime incident from the first to the second interview may be indicative of respondents purposively mis-reporting to avoid re-reporting a sensitive incident. However, the data cited above on self-reported reasons for changing suggests that a more dominant reason is respondents believe they understand the intent of the question better or re-considered whether their original answer best fit in with the intent of the survey. For example, the respondent may have thought the incident was within the reference period at the first interview, but re-dated it as occurring outside the reference period at the second interview. The fact that the changes are symmetrical in both directions is consistent with this idea. That is, reconsidering answers at the first interview lead to not only shifting a “yes” response at time 1 to “no” at time 2, but also vice versa.

The other changes at the person level are shifts from being a victim of rape or sexual assault at one interview to other unwanted sexual contact at the other interview. These respondents reported an incident with some type of unwanted behavior at both interviews, but did not report either force or inability to consent on the DIF at one of the interviews. For those who were classified as a rape victim, the changes are symmetrical across interviews (e.g., 9 shifted from time 1 to time 2 vs. 7 shifting in the other direction). For sexual assault there was a slightly larger shift to

⁴⁸“Not enough information” includes respondents who report an incident on the screener, but do not provide enough information on the DIF to classify the incident. For example, they refused to answer key items in the classification algorithm.

other unwanted behavior at the second interview (17) when compared to the shift in the opposite direction (n=11).

There are several reasons why a respondent shifts categories between interviews. It could be due to respondents not reporting the same incident at each time or it could be reporting the same incident but answering the questions on the DIF differently. The next section provides data related to the latter issue by focusing on the classification when the respondent reported the same incident.

Table 12-14. Estimates of reliability for detailed incident form classifications overall and by mode among the general population

Classification	Time 1=Yes ^a		Time 1=No ^a		Kappa ^b
	Time 2=Yes ^a	Time 2=No ^a	Time 2=Yes ^a	Time 2=No ^a	
Overall					
Rape and sexual assault ^c	9.8 %	6.7 %	4.0 %	79.6 %	0.58
Rape ^d	5.0	4.0	2.8	88.2	0.56
Completed ^e	2.4	2.3	2.7	92.6	0.46
Forced ^f	2.1	1.8	2.4	93.7	0.48
Unable to consent ^g	0.2	0.6	0.6	98.6	s
Attempted or threatened ^h	2.0	2.5	1.3	94.2	0.49
Sexual assault ⁱ	4.3	4.7	3.2	87.9	0.48
Completed ^e	3.2	4.3	2.9	89.6	0.43
Attempted or threatened ^h	0.7	1.4	1.6	96.4	s
Other unwanted sexual contact ^j	12.1	7.4	4.9	75.7	0.59
No unwanted behavior	0.4	4.1	3.2	92.4	s
ACASI^k					
Rape and sexual assault ^c	9.4 %	5.9 %	2.5 %	82.2 %	0.65
Rape ^d	4.7	3.1	2.1	90.1	0.62
Completed ^e	2.1	2.0	2.7	93.2	0.46
Forced ^f	2.0	1.7	2.3	94.1	0.48
Unable to consent ^g	0.0	0.4	0.8	98.8	s
Attempted or threatened ^h	1.9	2.1	0.7	95.3	0.57
Sexual assault ⁱ	4.5	4.0	1.4	90.1	0.60
Completed ^e	3.7	3.7	1.1	91.5	0.58
Attempted or threatened ^h	1.0	1.0	0.9	97.1	0.52
Other unwanted sexual contact ^j	12.2	5.7	4.7	77.4	0.64
No unwanted behavior	0.2	4.4	3.9	91.5	s

**Table 12-14. Estimates of reliability for detailed incident form classifications overall and by mode among the general population
(continued)**

Classification	Time 1=Yes ^a		Time 1=No ^a		Kappa ^b
	Time 2=Yes ^a	Time 2=No ^a	Time 2=Yes ^a	Time 2=No ^a	
CATIⁱ					
Rape and sexual assault ^c	10.5 %	8.3 %	7.2 %	74.1 %	0.48
Rape ^d	5.5	5.9	4.2	84.4	0.46
Completed ^e	2.8	3.1	2.7	91.4	0.47
Forced ^f	2.3	2.1	2.6	93.0	0.47
Unable to consent ^g	0.5	1.0	0.3	98.2	s
Attempted or threatened ^h	2.2	3.3	2.6	92.0	s
Sexual assault ⁱ	3.8	5.9	7.1	83.2	s
Completed ^e	2.1	5.5	6.7	85.7	s
Attempted or threatened ^h	0.0	2.2	2.9	94.9	s
Other unwanted sexual contact ⁱ	11.9	10.7	5.3	72.1	0.50
No unwanted behavior	0.7	3.5	1.6	94.2	s

Note: Estimates are based on unweighted data. Standard errors for kappa statistics can be found in Appendix A.

s Data suppressed for disclosure reasons.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview. "Yes" indicates the respondent reported experiencing that type of incidents in that interview. "No" indicates the respondent did not report experiencing that type of incident in that interview.

^bKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^cIncludes penetrative and non-penetrative sexual contact using force or while unable to consent.

^dIncludes penetrative sexual contact using force or while unable to consent.

^eIncludes incidents where the offender completed the behavior.

^fIncludes holding or pinning, using a weapon or threatening to use a weapon, other physical attacks or threats of physical attacks on respondent or someone else.

^gIncludes incidents where respondents were passed out for all or parts of the incident or were unable to consent due to alcohol or drugs.

^hIncludes incidents where the offender either verbal threatened or physically tried, but did not complete the behavior.

ⁱIncludes non-penetrative sexual contact using force or while unable to consent.

^jIncludes unwanted penetrative or non-penetrative sexual contact where force or while unable to consent was not reported. while unable to consent

^kAudio computer-assisted self-interview.

^lComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 12-15. Agreement rates and kappa statistic for respondents reporting victimization in the detailed incident form, for females in the general population sample, 2014-2015

Time 1 ^a	Time 2 ^a								Time 2 ^a agreement rate ^h	
	Past 12 months incidents ^b				Other unwanted sexual contact ^e	No unwanted behavior	Lifetime ^f only or not enough info ^g	No incident reported		
	All	Rape ^c	Sexual assault ^d							
Overall kappa ⁱ	0.58									
All respondents	870	69	49	113	12	557	70			
Past 12 month incidents ^b										
Rape ^c	72	44	5	9	1	11	2	61.1 %		
Sexual assault ^d	55	5	25	17	1	7	0	45.5 %		
Other unwanted sexual contact ^e	127	7	11	75	4	15	15	59.1 %		
No unwanted behavior	17	1	0	2	0	6	8	0.0 %		
Lifetime ^f incidents only or not enough info ^g	599	12	8	10	6	518	45	86.5 %		
Time 1 agreement rate ^g		63.8 %	51.0 %	66.4 %	0.0 %	93.0 %	0.0 %	76.1 %		

Note: Counts and agreement rates are based on unweighted data. Kappa estimation is based on weighted data. Standard errors for kappa statistics can be found in Appendix A.

^aTime 1 indicates the first interview. Time 2 indicates the re-interview.

^bIncludes incidents that occurred within the past 12 months of the interview.

^cIncludes penetrative sexual contact using force or while unable to consent.

^dIncludes non-penetrative sexual contact using force or while unable to consent.

^eIncludes penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

^fIncludes incidents that occurred at any point in the respondent's lifetime.

^gIncludes incidents where respondent said "no," refused to answer, or didn't know the answer to all behavior and tactic items.

^hIndicates percent of Time 2 reports that are the same as Time 1 reports.

ⁱKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance. For the purposes of kappa calculation, respondents in the Time 2 'No incident reported' group were combined with the Time 2 'Lifetime only or not enough info' group.

^jIndicates percent of Time 1 reports that are the same as Time 2 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Reliability of the Algorithm for the Same Incidents. To further explore the source of the changes observed in classification, the data were examined at an incident level using the reports that were matched using the narratives. The data presented in this section are unweighted and combine all of the re-interviews conducted with the GP and VO samples.

The overall kappa for the seven category crime typology is 48 percent, with a 63 percent overall agreement rate (table 12-16). Completed rape has relatively high agreement between the two interviews. Of the first interviews that are classified as completed rape, 74 percent are also classified the same way at the second interview. The shift into and out of completed rape between interviews is symmetrical, with both interviews classifying 43 incidents in this category. Many of the changes are related to reporting different unwanted behaviors at each interview. Of the 11 completed rapes reported at the first interview which changed categories at the second interview, 6 changed because the type of behavior is described as a sexual assault or attempted rape at the second interview. This compares to 2 that changed because the report of the tactic was different (see “other unwanted behavior”). This is consistent with the analysis above for the unwanted behavior questions (see table 12-5) in which a report of the unwanted behavior is not consistent between interviews. For example, one incident was described by a respondent the following way:

“....he started holding me down and trying to put his hand up my dress. I kept saying no that I didn’t want to, but he would just press down on me harder and flipped my dress up.I blacked out for a little bit and next thing I know we’re having sex.”

At the first interview, this respondent reported attempted penetration, whereas at the second interview completed penetration. There were a number of other instances like this, where the narrative provided by the respondent clearly describes the behavior, but the response to the behavior question at one of the interviews conflicts with the narrative.

The incidents involving attempted and threatened penetration had lower agreement rates than for completed penetration (e.g., Time 1 agreement 44% and 17% for attempted and threatened, respectively). The inconsistencies for these include both changing from attempt/threat to another type of rape (e.g., to completes) as well as changing from penetration to sexual assault. The latter

occurs because only non-penetrative sexual contact or non-contact was reported. For example, in one instance, the incident was described as follows

“He pushed me on the bed at some point and was on top of me and took his pants off but never exposed himself.He was just kissing on me and I kept telling him no.Then I pushed him really, really hard off of me and told him he had to leave. He finally left.”

In the first interview the respondent reported an attempt at penetration, whereas at the second interview, she did not report any type of unwanted attempt at penetration. While on its face, this incident involves a forced attempt at penetration, the respondent seemed to change her mind at the re-interview. An example of a more ambiguous situation follows:

“....he effectively came at me and started trying to unbutton my shirt and was holding me down with his body weight and I told him to stop and he got back to his senses and stopped....”

In this instance, the respondent reported on the DIF a forcible attempt at penetration at the first interview and no attempt at penetration at the second interview. There were several other examples where the respondent changes her interpretation of the perpetrator’s intent.

Sexual assault has much lower agreement rates than for rape, with 49 percent of the completed sexual assaults at time 1 being classified in the same category at time 2. Attempted sexual assault is lower than this. Many of these inconsistencies was movement between interviews for incidents classified as sexual assaults at one interview and other unwanted contact at the other interview. This shift in classification occurred because the respondent reported the tactic as “force” at one interview and something else at the other interview (e.g., no tactic at all; coercion). For example of those who reported completed sexual assault at the first interview (n=65), about one-third were classified as other unwanted contact (n=20) at the second interview. A similar pattern occurs when comparing the completed sexual assaults at the second interview to the first interview, where 10 were classified as other unwanted contact. A review of the narratives of these incidents reveals that most are instances where the victim was groped in public or unexpectedly, such as illustrated below:

“....As we were talking to the guy sitting next to us, he made a rather rude comment and then reached out and grabbed my breast.....”

“....walking home from a friends dorm, group of men pass me on the street. one of them lifts my skirt up and grabbed my ass as I walked by, and continued in the opposite direction as me.”

For these types of incidents, the tactics listed on the questionnaire did not fit the situation. The survey included physical attacks or threats, such as pinning down the victim, but did not include an option such as “grabbed or groped.” The respondent could write in this tactic in an “other specified” field, which is how these were picked up at one of the interviews. The inconsistencies noted above are cases where the respondent wrote in the tactic on the questionnaire at one interview, but did not write it in at the other interview.

Most discrepancies in classification between interviews occurred for the attempted or threatened sexual assaults. Unlike the other types of crimes, there was not a symmetrical change in the number of incidents reported at each interview. At the first interview there were 27 incidents classified as an attempted assault while there were only 12 incidents classified as this or a threat at the second interview. Much of the difference between the two interviews involved incidents switching from an attempt/threatened sexual assault to other unwanted behavior at the second interview. For example, of the 20 incidents which changed from attempted sexual assault at the first interview to something else at the second interview, 12 were classified as other unwanted contact. This indicates neither force or inability to consent was reported at the second interview. Several of these 12 were similar to the above where the respondent reported the tactic (grabbing, groping) as an “other specified” at for the first interview, but did not do so at the second interview. However, there were also a number of instances where the event itself was not clear-cut. These involved unexpected advances that were stopped after the first or subsequent attempts:

“...There were only three of us left and there were-- two of us were laying down and the third person went in for a kiss. And I basically just moved away, turned around, and it stopped there.”

“. . .He tried to kiss me three times and pull me over to him. After I shouted at him, he stopped.”

In both of these instances the respondent reported the force at the first interview, but not at the second interview.

Table 12-16. Agreement rates and kappa statistic for detailed incident form classification of crime type at the first and second interviews, 2014–2015

Time 1 ^a	All incidents	Time 2 ^a					
		Rape ^b			Sexual assault ^c		
		Completed ^d	Attempted ^e	Threatened ^f	Completed ^d	Attempted ^e	Threatened ^f
Overall kappa ⁱ	0.48						
All incidents	341	43	19	9	59	11	1
Rape ^b							
Completed ^d	43	32	2	0	4	0	0
Attempted ^e	18	3	8	2	3	0	0
Threatened ^f	12	1	3	2	1	0	0
Sexual assault ^c							
Completed ^d	65	3	5	1	32	3	0
Attempted ^e	27	0	1	0	5	7	0
Other unwanted sexual contact ^f	149	1	0	3	10	1	1
Not enough information	27	3	0	1	4	0	0
Time 2 ^a agreement rate ^j		74.4 %	42.1 %	22.2 %	54.2 %	63.6 %	0.0 %

Table 12-16. Agreement rates and kappa statistic for detailed incident form classification of crime type at the first and second interviews, 2014–2015 (continued)

Time 1 ^a	Time 2 ^a			Time 1 ^a agreement rate ^h
	Other unwanted sexual contact ^g	Not enough information		
Overall kappa ⁱ				
All incidents	175	24		
Rape ^b				
Completed ^d	3	2	74.4 %	
Attempted ^e	2	0	44.4	
Threatened ^f	5	0	16.7	
Sexual assault ^c				
Completed ^d	20	1	49.2 %	
Attempted ^e	12	2	25.9	
Other unwanted sexual contact ^g	124	9	83.2 %	
Not enough information	9	10	37.0 %	
Time 2 ^a agreement rate ^j	70.9 %	41.7 %	63.0 %	

Note: Estimates are based on unweighted data.

^aTime 1 indicates the first interview. Time 2 indicates the reinterview.

^bIncludes penetrative sexual contact using force or while unable to consent.

^cIncludes non-penetrative sexual contact using force or while unable to consent.

^dIncludes incidents where the offender completed the behavior.

^eIncludes incidents where the offender physically tried, but did not complete the behavior.

^fIncludes incidents where the offender verbally threatened, but did not physically attempt the behavior.

^gIncludes unwanted penetrative or non-penetrative sexual contact where force or while unable to consent was not reported.

^hIndicates percent of Time 1 reports that are the same as Time 2 reports.

ⁱKappa is a measure of agreement between interviews. It is a more robust measure than percent agreement, as it takes into account agreement occurring by chance.

^jIndicates percent of Time 2 reports that are the same as Time 1 reports.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

12.4 Summary and Discussion

The screener items, when grouped into logical categories, have reliabilities of 60 to 70 percent, depending on which mode is examined; these reliabilities are considered to be “substantial” using the Koch and Landis (1977) standard. Nonetheless, this seems relatively low, given the fairly specific behaviors that are being referenced on the screener. A significant number of respondents are changing their reports between interviews. For example, approximately 32 percent of respondents who reported a completed rape to the screener items at the first interview changed at the second interview. The change seemed to be distributed across the remaining types of victimizations, including not reporting any incident at all. The sample sizes for the general population sample were small for rape incidents, but the result was consistent for the volunteer sample as well. When examined at a person level, the screener items that had the lowest consistency were those related to other unwanted completed or attempted penetration. Slightly more than half of respondents changed how they answered these screener items. The most consistent patterns were those who only reported a lifetime incident. For example, 84 percent of those reporting only a lifetime incident at the first interview reported the same incident at the second interview. Most of the 16 percent who changed at the reinterview shifted to reporting no incidents at all.

When respondents were asked why they changed their answers, the most common reasons related to some type of respondent conditioning in the form of understanding the questions differently or thinking further about their experiences between interviews. To some degree, this type of conditioning reflects the complex structure of the BSQs, where most include three different conditions (i.e., behavior, tactic, and consent). The cognitive interviews conducted for the development of these items found that some respondents may not consider all three of the conditions (Steiger et al., 2014).

For the 12-month items, some also reported not responding to the 12-month items to avoid being asked the follow-up questions. A pattern consistent with this was observed for individuals who reported an incident occurring during the 12-month reference period at the first interview but not at the second interview. However, the empirical patterns noted above indicate that this is only one of several contributors to the inconsistencies between interviews. For example, of the 32 percent of those changing from reporting a rape on the screener at the first interview to something else at the second interview, a relatively small percentage (7 of the 32 percent) did not report any type of incident. Other reasons for the changes, such as rethinking which screener items may be most appropriate, were among the more common reasons respondents reported.

The inconsistencies related to the DIF items reflect several issues. The least reliable DIF items in the classification algorithm were the measures of unwanted behaviors. The measures of attempted and threatened behaviors were the most likely to change between the two interviews. When answers changed, they went from an attempt or threat to not reporting the occurrence of any unwanted behavior. To a lesser degree, there were also inconsistencies related to reporting a completed behavior. This is consistent with the comparison of the screener and DIF classifications (see Chapter 9). The questions on unwanted behaviors, which asked respondents to make distinctions between completed, attempted, and threatened acts.

The items used to classify the incident as using force and inability to consent had relatively high agreement rates and reliabilities. The two main items—the type of force such as holding or pinning the victim down and saying she was unable to give consent—had agreement rates close to 90 percent. For assaults, error was found in the question on force, which omitted a category for groping or grabbing. A significant number of respondents wrote in this tactic, but some did not. This led to changes of incidents classified as sexual assault at one time and other unwanted behavior at another time.

Finally, some of the inconsistencies related to the DIF items reflect ambiguities of the incident itself. Incidents that involve persistent, and aggressive, advances were not consistently classified as using force by the respondent.

Overall, this analysis indicates that much of the inconsistency in final classification can be addressed by redesign of several key items on the detailed incident form, including improving the questions on completed vs. attempted/threatened acts and expanding the list of tactics to include groping and grabbing.

There are several limitations of the above analysis; the most important are the sample sizes available for analysis. Once narrowing the focus to specific types of incidents, the number of respondents that endorsed particular items is not large. At the screener, the common pattern across both modes and types of samples reinforces the main conclusions. For the DIF, the responses were combined, which limited the possibility of testing for differences across modes. This also narrowed the focus to those incidents where respondents provided narratives at both times. Additional analyses not reported above found that respondents were more likely to provide a narrative in the CATI mode, and were in the volunteer sample, for incidents not involving completed rape and if they reported one or two (as opposed to three) 12-month incidents.

13. Cost Effectiveness: Comparisons of ACASI and CATI Designs

This chapter discusses the relative cost of the ACASI and CATI interviewing methods. The first section provides estimates of the relative costs of the two methods. The second section provides information on the relative costs of the landline and cell phone portions of the CATI interviews.

13.1 Relative Cost Analysis

This section examines the relative cost of data collection using telephone survey administration (computer-assisted telephone interviews, a.k.a. CATTI) and household in-person administration (audio computer-assisted self interviews, a.k.a. ACASI). The cost-per-completed-survey was calculated to compare the two modes and presented in this chapter as ratios. This aggregation takes into account the different types and quantity of sample and the different number of cases by mode as shown in table 13-1.

Table 13-1. Completed surveys by mode and sample type

	Main survey	Re-interview	Total
In-person completed surveys			
ABS sample	3,053	359	3,412
VO sample	1,012	154	1,166
SP sample	41	-	41
Total	4,106	513	4,619
Telephone completed surveys			
Landline sample	965	33	998
Cell phone sample	4,222	469	4,691
VO sample	1,162	102	1,264
SP sample	17	-	17
Total	6,366	604	6,970

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Across all sample types, in-person data collection costs are 4.1 times higher than for telephone data collection. Cost components for both modes were broken out by interviewer training, field/telephone interviewing, and respondent incentives. For both in-person and telephone

modes, field/telephone interviewing was the largest expense category, representing 75 percent of the total costs as shown in table 13-2 below.

The ratio was somewhat higher for the volunteer sample. This is a slight over-estimate of the costs since it does not account for the screening that was done for the in-person survey to find households with women 18-49 years old.

Table 13-2. Cost components by mode

	Percent of total cost	Ratio to telephone mode
In-person completed surveys		
Training	21%	5.8
Data collection	75	4.0
Respondent Incentives	4	2.1
Total	100%	4.1
Telephone completed surveys		
Training	15%	
Data collection	77	
Respondent Incentives	8	
Total	100%	

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Although both modes required screening to identify eligible respondents (ABS sample for in-person and RDD landline sample for telephone), in-person screening involved substantially higher amounts of labor and travel to accomplish. In-person screening costs were minimized through the use of a roster mail survey conducted prior to the data collection period. Households that returned the roster made it possible to identify ineligible households rather than dispatching field interviewers to the sampled addresses. This reduced the number of cases sent to the field by 19 percent and reduced field costs by up to 15 percent. A complete description of the roster mailing to ABS addresses and the method for calculating the cost savings is provided in Appendix J.

Field/telephone interviewing costs include all labor, non-labor, and indirect costs incurred during the data collection periods for the tasks listed below:

- household mail roster mailing, receipt, and processing (in-person ABS)
- pre-contact letter mailing (RDD landline sample matched to mailing addresses)
- interviewer attempts to reach sampled addresses, householders, cell phone users, and respondents
- field and telephone supervision, data collection management, and home office support

- coaching, retraining, and monitoring for performance, productivity, and distress
- travel to/from sampled addresses and scheduled interview appointments (in-person mode only).

Interviewer training was the next largest expense category for both modes, constituting 21 percent of the in-person data collection costs and 15 percent of the telephone costs. In-person training was 5.8 times more expensive than the remote training conducted for telephone interviewers. The primary cost drivers for the training were the direct expenses (travel, lodging, meals, and venue rental for in-person vs. web-based distance learning platforms for telephone) and trainee labor hours (30 hours for in-person vs. 16 hours for telephone).

In-person training costs were minimized through the use of online learning modules assigned to all field trainees and completed prior to the start of training. Sessions were held in each of the five CBSAs at centrally located venues to maximize the number of trainees who could attend as daily commuters. This reduced the number of trainees attending as travelers and minimized the cost of lodging and per diem.

Interviewer training costs include all labor, non-labor, and indirect costs incurred during the interviewer training periods for the tasks listed below:

- interviewer staff hiring and training
- development of field procedures, administrative forms, and materials
- development of procedure manuals, curricula, and training sessions
- logistics planning and procurement of training venues
- travel expenses (in-person mode only)
- web-based instruction and on-line sessions (telephone mode only).

Respondent incentives were 8 percent of the telephone data collection costs and 4 percent of the in-person costs. The payment structure was different for the two modes. All telephone respondents who completed the main or re-interview surveys received a \$20 check. The cost of these incentives reflected only those respondents who provided an address to receive the incentive check and those who cashed their check. Approximately 10 percent of the respondents did not provide an address and 7 percent of the mailed checks were never cashed. The in-person respondents who completed the main or re-interview surveys received differential incentives by

sample type. ABS sample respondents received \$20, VO and Service Provider (SP) respondents received \$30, and SP respondents were offered an additional \$10 (\$40 total) to cover the cost of any required travel (e.g., if the interview was conducted at the service provider agency). All in-person incentives were provided via a debit card. Unlike the checks mailed to telephone respondents, the value of the debit card was charged in full regardless of whether or not the respondent used it.

13.2 Relative Cost by RDD Sample Type

Diminishing response rates reflect the difficulty in reaching respondents through RDD samples due to their reluctance to answer phones from unknown callers. This section examines the relative cost of the two RDD sample types—landline and cell phone numbers. Differences between these samples affect the relative cost of telephone data collection. Screening for eligible respondents (female, age 18 or older) was conducted for landline numbers at the household level as the landline number is usually shared by all members of the household. Screening of cell phone numbers was done at the individual level because cell phones are used as individual devices and rarely shared. Although both sample types had relatively low screening response rates (21.3% landline, 32.5% cell phone),⁴⁹ greater effort was required to complete eligibility screeners of the landline sample than cell phones. Landline households were less likely to answer the phone, but more likely to have an eligible respondent as a resident. Cell phone users were more likely to answer the phone, but over half were screened out due to sex (not female) or age (younger than 18 years old).

Both RDD types required large sample sizes (21,312 landline phone numbers, 105,498 cell phone numbers) and a high volume of call attempts (235,388 landline dials, 882,521 cell phone dials) to reach completion targets. On average, landline numbers required approximately 32 percent more call attempts than cell phone numbers with 11.0 dials per number versus 8.4 cell phone dials. Comparing the average number of call attempts per completed interview provides a stronger driver of relative cost because the screener and survey interviews involve interviewer labor. Landline numbers required 236 call attempts and cell phone numbers required 188 call attempts per complete. This shows that the cell phone sample had a 25 percent efficiency advantage over the landline sample. Another cost measure is the quantity of sampled telephone numbers required to complete the interviews. These differences are presented in table 13-3 below.

⁴⁹See Chapter 5, Response Rates, tables 5-2 and 5-3.

Table 13-3. RDD productivity comparison

RDD sample	Landline	Cell phone	Total
Sampled telephone numbers	21,312	105,498	126,810
Call attempts	235,388	882,521	1,117,909
Dials per attempt	9.2	7.9	
Completed interviews	998	4,691	5,689
Call attempts	235,388	882,521	1,117,909
Dials per complete	235.9	188.1	
Sampled telephone numbers	21,312	105,498	126,810
Completed interviews	998	4,691	5,689
Sample per complete	21.4	22.5	22.3

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

A key measure for this study is the incidence and prevalence of sexual assault for women in the past 12 months. Women in the 18-29 age category reported the most assaults in the past year, followed by those in the 30-39 age category. Very few sexual assaults in the past 12 months were reported by respondents older than 40. Examining the yield from the telephone sample to complete a main interview broken out by respondent age category provides the strongest measure of relative cost. The yield of completed interviews by the landline versus cell phone sample is most pronounced for the age 18-29 and 30-39 categories. Only 5 percent of the completed main interviews from the landline sample were completed by respondents ages 18-29. However, the cell phone sample produced 24 percent of the main interview completes from this age group. Stated another way, the 21,312 landline sample produced only 51 main interview completes from women age 18-29, or 418 landline numbers per complete. The 105,498 cell phone sample produced 1,010 main interviews from women age 18-29, or 104 cell phone numbers per complete. This illustrates that the amount of landline sample required was 4.0 times greater than the cell phone sample to yield a completed main interview by an age 18-29 respondent. The greater efficiency of the cell phone sample was also evident for the 30-39 and 40-49 age groups, but to a lesser level—2.6 times more landline sample was required to yield completes from 30- to 39-year-old respondents and 1.2 times more for 40- to 49-year-old respondents. It is only for the age 50 or older category that the landline sample produced a higher efficiency than cell phone by a factor of 2.1. Details for these efficiency measures are shown in table 13-4 below.

Table 13-4. RDD productivity by age group

RDD sample	Landline		Cell phone		Total
Sampled telephone numbers	21,312		105,498		126,810
Main interviews by age category	Cases	%	Cases	%	Total
18-29	51	5%	1010	24%	1061
30-39	58	6%	749	18%	807
40-49	120	13%	733	17%	853
50 or older	731	76%	1719	41%	2450
	960		4211		5171
Sample used by age					
18-29	418		104		522
30-39	367		141		508
40-49	178		144		322
50 or older	29		61		91

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

These data demonstrate that the cell phone sample had higher productivity than the landline sample, both in terms of the number of call attempts required to produce a completed interview and the amount of sample required to yield a completed interview from respondents in the age category who are most at risk of sexual assault.

14. Respondent Reactions and Interviewing Environment

The RSA Pilot Test put in place a number of procedures to mitigate the risks that respondents may experience when taking the survey. One is the risk of becoming emotionally upset by asking about sexual violence. For those that have experienced an incident, the act of remembering and talking about the incident may evoke emotions that would normally not be part of a survey interview. Research to date has found that, in fact, surveys on sexual violence do evoke these types of emotions, but respondents generally work through these feelings and express an overall positive experience doing the interview (DePrince & Chu, 2008; McClinton, Lund, de Vries, & Matthews, 2015; Newman, Risch, & Kassam-Adams, 2006). A second type of reaction may occur for those that find the language on the sexual victimization screener to be too graphic. The first section of this chapter presents results of a series of debriefing questions that asked respondents about their feelings, as well as their overall experience related to taking the survey. The second section summarizes interviewer observations related to the survey conditions.

14.1 Respondent Reactions to the Survey

At the end of the survey, respondents were asked a series of debriefing questions about their experience participating in the study. These questions were a series of items using a five-point agree-disagree scale (strongly agree, agree, neutral, disagree, strongly disagree; see items DQ1 – DQ7 in Appendix B). These were administered in both the ACASI and CATI modes. These questions asked about potentially negative reactions, such as experiencing intense emotions, thinking about things they didn't want to think about, and raising unpleasant emotions that were not expected. The debriefing continued with questions about positive reactions, such as believing they had helped others by participating, gaining something positive from the study, and being glad to have had the opportunity to participate. The debriefing concluded by asking how easy or hard the questions were to understand and whether they would have made the same choice to participate now that they know what the survey is about.

14.1.1 Emotional Reactions to the Survey

Generally, more respondents reported higher agreement with items indicating positive reactions to the survey than negative reactions, and most respondents indicated that they did not regret taking the survey (tables 14-1 and 14-2). Among the general population (GP) sample, significantly more respondents in the CATI mode agreed or strongly agreed that they experienced both positive and negative reactions to the survey than respondents in the ACASI mode (table 14-1). CATI respondents were also more likely than ACASI respondents to say they would take the survey again now that they know what it is about (i.e., do not regret taking the survey) (90.1% vs. 78%, respectively). The pattern among the volunteer (VO) sample, however, was somewhat different (table 14-2). Among the VO sample, significantly more respondents in the ACASI mode than in the CATI mode agreed or strongly agreed that they experienced negative reactions to the survey. CATI respondents were somewhat more likely to endorse positive reactions to the survey than ACASI respondents, and were significantly more likely to agree they did not regret taking the survey, although the difference is not substantively meaningful (94.9% vs. 91.9%, respectively).

Table 14-1. Respondents' reactions to the survey, by mode of interview for females ages 18-49 in the general population, 2014-2015

	Total	
	ACASI^a	CATI^{b*}
Negative reactions		
You experienced intense emotions	22.2 %†	26.0 %
Made you think about things you didn't want to	30.6 †	35.6
Raised unpleasant emotional issues that you had not expected	17.3 †	21.1
Positive reactions		
Believe you have helped others by participating	67.9 %†	82.4 %
You were glad to have the opportunity to participate	75.0 †	81.7
You gained something positive from participating	54.1 †	65.4
Lack of regret		
You would make the same choice to participate	78.0 †	90.1
Number of weighted sample cases	11,178,108	10,556,211

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-2. Respondents' reactions to the survey, by mode of interview for females ages 18-29 in the volunteer sample, 2014-2015

	Total	ACASI^a	CATI^{b*}
Negative reactions			
You experienced intense emotions	29.9 %†	21.0 %	
Made you think about things you didn't want to	45.7 †	37.4	
Raised unpleasant emotional issues that you had not expected	24.7	22.3	
Positive reactions			
Believe you have helped others by participating	81.1 %†	86.8 %	
You were glad to have the opportunity to participate	94.3	92.9	
You gained something positive from participating	74.9	74.7	
Lack of regret			
You would make the same choice to participate	91.9 †	94.9	
Number of unweighted sample cases	981	1,087	

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

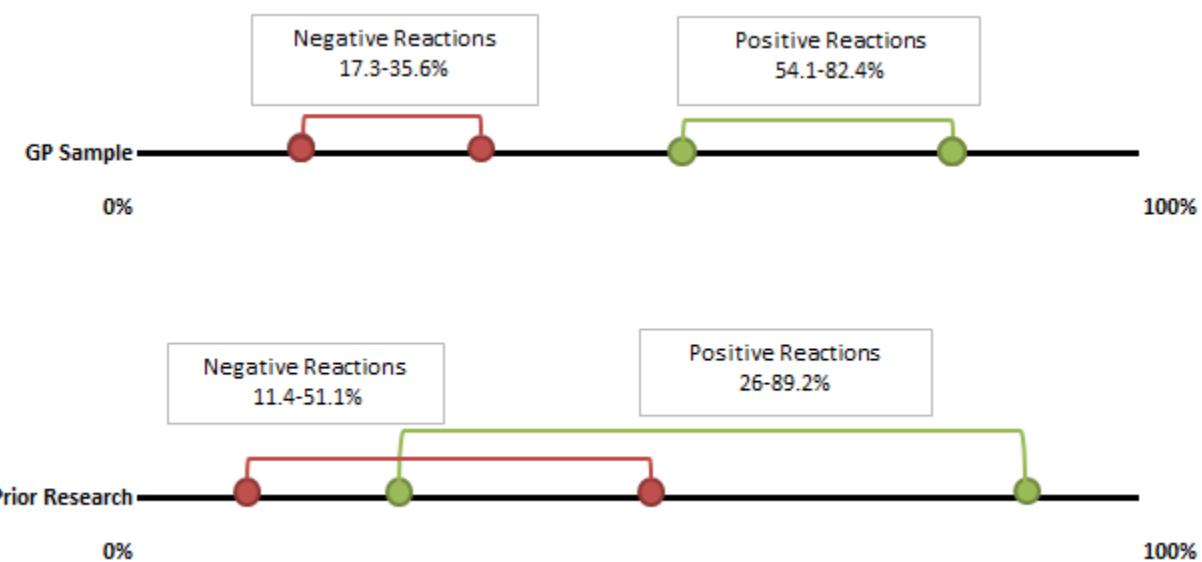
^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

14.1.2 Comparisons to Other Studies

In general, our findings are consistent with prior research that has found sensitive surveys on topics such as sexual victimization tend to produce more self-reported positive reactions than negative reactions and that most respondents do not regret taking such surveys (e.g., DePrince & Chu, 2008; McClinton et al., 2015; Newman et al., 2006). Although it is difficult to compare rates of specific respondent reactions among studies as they use different methodologies, samples, and questions, our rates do fall within ranges reported elsewhere. The most comparable sample in the current study to prior research are the GP respondents. Among the GP sample, endorsement of negative reactions ranged from 17.3 to 35.6 percent and endorsement of positive reactions ranged from 54.1 to 82.4 percent (figure 14-1). Prior studies that measured respondent reactions to sensitive surveys with somewhat comparable research designs report negative reactions ranging from 11.4 to 51.1 percent and positive reactions ranging from 26 to 89.2 percent (Black et al 2006; Valpied et al, 2014; Wager, 2012; Walker et al, 1997). Regret has been measured less often, but Wager (2012) and Walker et al. (1997) included items similar to the RSA Pilot Test, with 75.8 to 84.3 percent respondents agreeing or strongly agreeing that they did not regret taking the survey. GP respondents reported similarly high levels of lack of regret, ranging from 78 to 90.1 percent.

Figure 14-1. Range of respondents in the GP sample and prior research studies that reported negative and positive reactions



14.1.3 Victimization

This section examines the relationship between emotional reactions to the survey and reporting a victimization.

Recency of Victimization

Respondents in the GP sample who had been victimized either in the last 12 months or in their lifetime reported higher agreement with items indicating both positive and negative reactions to the survey than non-victims (table 14-3). For example, 20.3 percent of respondents with no victimizations, 27.9 percent of respondents with lifetime-only victimizations, and 33.3 percent of respondents with victimizations within the last year said that they experienced intense emotions. Showing a similar pattern, 70.7 percent of non-victims, 80.4 percent of lifetime victims, and 81.7 percent of victims within the last 12 months believed they had helped others by participating. Non-victims in the VO sample (table 14-4) also reported less negative reactions to the survey than both types of victims, but they did not tend to differ from victim responses regarding positive reactions.⁵⁰

⁵⁰These general findings—that non-victims in both sample types report less negative reactions and that non-victims in the GP sample also report less positive reactions—appear to hold regardless of mode.

Table 14-3. Respondents' reactions to the survey, by type of victimization for females ages 18-49 in the general population, 2014-2015

	Total	Past 12 months only	Lifetime only	No victimizations*
Negative reactions				
You experienced intense emotions	24.0 %	33.3 %†	27.9 %†	20.3 %
Made you think about things you didn't want to	33.0	54.2 †	43.2 †	23.6
Raised unpleasant emotional issues that you had not expected	19.2	33.7 †	25.7 †	12.9
Positive reactions				
Believe you have helped others by participating	74.9 %	81.7 %†	80.4 %†	70.7 %
You were glad to have the opportunity to participate	78.3	82.5 †	81.6 †	75.7
You gained something positive from participating	59.6	64.3	57.1	60.1
You would make the same choice to participate	83.9	84.4	88.2 †	81.5
Ease of understanding				
How easy or hard were the questions to understand	80.2 %	73.2 %†	78.2 %†	82.5 %
Number of weighted sample cases	21,738,717	2,371,337	6,750,785	12,637,170

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-4. Respondents' reactions to the survey, by type of victimization for females ages 18-29 in the volunteer sample, 2014-2015

	Total	Past 12 months only	Lifetime only	No victimizations*
Negative reactions				
You experienced intense emotions	25.2 %	34.5 %†	25.0 %†	13.1 %
Made you think about things you didn't want to	41.3	54.2 †	44.1 †	21.2
Raised unpleasant emotional issues that you had not expected	23.4	32.8 †	24.1 †	10.3
Positive reactions				
Believe you have helped others by participating	84.1 %	83.7 %	87.0 %†	81.5 %
You were glad to have the opportunity to participate	93.6	92.4	95.3	93.3
You gained something positive from participating	74.8	75.2	73.6	75.5
You would make the same choice to participate	93.5	90.3 †	96.1	95.1
Ease of understanding				
How easy or hard were the questions to understand	84.2 %	78.0 %†	86.4 %†	90.0 %
Number of unweighted sample cases	2,068	815	640	613

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

As described previously, VO respondents in general do not show as many mode differences as GP respondents, so it is possible other factors affect respondent reactions for that sample. VO respondents initiated contact with the survey and were, by definition, more cooperative than GP respondents, who were asked to participate. This may have been more important to reporting their reactions to the survey than their victimization history.

While negative reactions are lowest for those that did not report a victimization in both GP and VO samples, there are still a nontrivial number of respondents who reported some type of negative reaction. This may be related to the explicit language in the questionnaire or to the general negative emotions associated with the topics that are being covered on the survey. Our finding that victims were more likely to experience negative emotions is consistent with prior research (e.g., Decker et al., 2011; DiLillo et al., 2006; Jaffe et al., 2015; McClinton et al., 2015; Wager, 2012; Walker et al., 1997). Extant research on survey benefits is less common, but some evidence also supports our findings that victims were more likely to endorse positive reactions (Decker et al., 2011; Wager, 2012).

Number of DIFs Completed

We also examined differences among respondents who had experienced victimization by comparing responses from those who completed one DIF, two DIFs, and three or more DIFs. In general, respondents who completed one DIF were less likely to report negative reactions to the survey than those who completed three or more DIFs in both GP and VO sample types (tables 14-5 and 14-6). Reactions to the survey did not differ consistently between respondents who completed one versus two DIFs, and the number of DIFs completed did not generally affect respondents' positive reactions or regret for taking the survey.

Table 14-5. Respondents' reactions to the survey, by number of detailed incident forms completed for females ages 18-49 in the general population, 2014-2015

	One detailed incident form completed	Two detailed incident forms completed	Three or more detailed incident forms completed*
Negative reactions			
You experienced intense emotions	26.6 %†	41.2 %	38.0 %
Made you think about things you didn't want to	47.5 †	54.4	63.0
Raised unpleasant emotional issues that you had not expected	29.1	37.4	38.1
Positive reactions			
Believe you have helped others by participating	80.8 %	81.9 %	83.3 %
You were glad to have the opportunity to participate	80.3	80.9	87.1
You gained something positive from participating	65.3	60.3	65.7
You would make the same choice to participate	86.4	74.4 †	87.1
Ease of understanding			
How easy or hard were the questions to understand	74.3 %	71.6 %	72.9 %
Total number of respondents	1,119,458	464,113	783,226

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-6. Respondents' reactions to the survey, by number of detailed incident forms completed for females ages 18-29 in the volunteer sample, 2014-2015

	One detailed incident form completed	Two detailed incident forms completed	Three or more detailed incident forms completed*
Negative reactions			
You experienced intense emotions	28.6 %†	32.9 %	40.0 %
Made you think about things you didn't want to	43.2 †	51.9 †	64.1
Raised unpleasant emotional issues that you had not expected	30.0 †	27.8 †	37.3
Positive reactions			
Believe you have helped others by participating	81.9 %	87.2 %	83.7 %
You were glad to have the opportunity to participate	90.9	95.6	92.1
You gained something positive from participating	73.5	79.1	74.9
You would make the same choice to participate	92.0	93.7 †	87.6
Ease of understanding			
How easy or hard were the questions to understand	84.0 %†	77.2 %	74.1 %
Total number of respondents	287	158	370

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

14.1.4 Survey Ease of Understanding

Most respondents found the survey questions easy to understand. Among the GP sample (table 14-7), significantly more respondents agreed that the survey was easy to understand if they answered questions in the ACASI mode (84.5%) than CATI mode (75.6%). This finding also held true for the VO sample (table 14-8).

Table 14-7. Respondents' ease of understanding the survey, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	Total		Past 12 months only	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
Questions were easy to understand	84.5 %†	75.6 %	77.9 %†	65.7 %
Number of weighted sample cases	11,178,108	10,556,211	1,459,509	913,491

Table 14-7 (continued)

	Lifetime only		No victimizations	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
Questions were easy to understand	86.2 %†	74.1 %	85.2 %†	78.6 %
Number of weighted sample cases	2,273,685	4,477,101	7,454,961	5,182,209

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-8. Respondents' ease of understanding the survey, by type of victimization and mode of interview for females ages 18-29 in the volunteer sample, 2014-2015

	Total		Past 12 months only	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
Questions were easy to understand	89.0 %†	79.9 %	84.9 %†	66.0 %
Number of unweighted sample cases	981	1,087	518	297

Table 14-8 (continued)

	Lifetime only		No victimizations	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
Questions were easy to understand	92.3 %†	84.1 %	94.3 %†	86.4 %
Number of unweighted sample cases	181	459	282	331

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Summary

Although precautionary approaches to mitigate the risk of harm to respondents are important, research shows that in general, respondents do not report substantial harm from taking sensitive surveys. Even those who do report some negative feelings do not typically regret taking the survey because they also report experiencing positive emotions such as empowerment or contributing to solving the problem (DePrince & Chu, 2008; McClinton et al., 2015; Newman et al., 2006). Findings from the current study are in line with this prior research, with rates for positive and negative reactions falling within ranges reported elsewhere (Black et al., 2006; Valpied et al., 2014; Wager, 2012; Walker et al., 1997). Respondents in the current study reported higher agreement with items indicating positive reactions than negative reactions, and most respondents indicated that they did not regret taking the survey. This general pattern held true across survey modes and levels of victimization, although both survey methodology and respondent characteristics were significantly related to respondent reactions. In sum, multiple measures combine to indicate that the risk of harm for respondents participating in this study was minimal. The substantial majority of respondents were able to balance negative and positive reactions, completed the survey without distress, and did not regret their participation.

14.2 Interviewer Observations

Each interviewer was asked to provide assessments of the interview. One such assessment was judging whether the respondent was displaying any signs of emotional distress at the end of the interview. In addition, interviewers were trained to complete an interviewer observations questionnaire (IOQ) as a part of data collection (table 14-9). The purpose was to document whether the survey was conducted in privacy, to note signs of distress, and to record other factors that could have influenced the way respondents chose to answer items on the survey. Methodological and environmental factors such as the presence of others during the interview may affect data quality (Coker & Stasny, 1994; Catalano, 2016). Therefore, it was important to assess how well the procedures were carried out and to examine if these had any effects on the respondents, given the sensitive nature of the survey.

Table 14-9. Interviewer observations questionnaire items by mode

	CATI	ACASI
Cooperation of the respondent	✓	✓
Verbal distress	✓	✓
Nonverbal distress	✓	✓
Listening in on any part of the interview	✓	
How much of the time someone listening	✓	
Interview setting		✓
Resistance to private setting		✓
Who resisted		✓
Asked questions during the ACASI interview		✓
Types of questions asked		✓
Respondent wore headphones during ACASI		✓
Respondent looked at Event History Calendar		✓
Anyone else in the room during interview		✓
Length of time person in the room		✓
Portion of the interview someone in the room		✓
During ACASI someone looking over shoulder		✓
Who was in the room		✓

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Particular IOQ items were tailored for each method of data collection and were asked regardless of mode. These included assessments of the respondent's level of cooperation and signs of verbal and nonverbal distress. Other items were specific to the setting or context and varied by mode. For those interviewed by phone via CATI, interviewers were asked to document if it seemed like someone was listening in on any part of the interview and to note the percentage of time. For those administered in person via ACASI, interviewers were asked to document the presence of someone else during any part of the interview, who was in the room and during what portion of the interview, and the behavior of the person in the room (i.e., walking/standing in the room, looking over the respondent's shoulder, etc.) (see Appendix B for a complete listing of items).

This section presents key findings of interviewer observations documented at the end of each survey administration.

14.2.1 Interviewer Ratings of Distress

To supplement respondent self-reported survey experiences, interviewers rated the respondent's level of distress, as observed by the interviewers just prior to the conclusion of the survey. In general, interviewers rated the overwhelming majority of respondents as neutral (i.e., showing no signs of distress). Only one CATI respondent broke off an interview due to distress. However, no respondents were rated the highest "elevated" distress level in either sample type. Among GP respondents, interviewers rated 97 percent of respondents in both the CATI and ACASI modes as neutral and 3 percent of respondents as exhibiting "low-to-moderate" distress (table 14-10). The percent of respondents showing "low-to-moderate" signs of distress rose with severity of victimization, as respondents who had been victimized within the last 12 months were most likely to be observed as showing these signs of distress (9.1% ACASI and 11.7% CATI modes) and non-victims were the least likely to be observed showing these signs of distress (1.8% ACASI and .6% CATI). Differences between modes were only significant for non-victims. This pattern was generally similar for respondents in the VO sample (table 14-11).

Table 14-10. Interviewer ratings of distress, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	Total		Past 12 months only	
	ACASI ^a	CATI ^b	ACASI ^a	CATI ^b
Interviewer assessment of respondent emotional state				
Neutral	97.0 %	97.1 %	90.9 %	88.3 %
Low to moderate	3.0	2.9	9.1	11.7
Elevated	--	--	--	--
Number of weighted sample cases	11,262,817	10,573,806	1,477,168	934,526

Table 14-10 (continued)

	Lifetime only		No victimizations	
	ACASI ^a	CATI ^b	ACASI ^a	CATI ^b
Interviewer assessment of respondent emotional state				
Neutral	96.9 %	96.3 %	98.2 %†	99.4 %
Low to moderate	3.1	3.7	1.8 †	0.6
Elevated	--	--	--	--
Number of weighted sample cases	2,273,685	4,461,654	7,511,965	5,177,626

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

-- Less than 0.05%.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-11. Interviewer ratings of distress, by type of victimization and mode of interview for females ages 18-29 in the volunteer sample, 2014-2015

	Total		Past 12 months only	
	ACASI ^a	CATI ^b	ACASI ^a	CATI ^b
Interviewer assessment of respondent emotional state				
Neutral	96.8 %†	98.7 %	95.2 %	95.8 %
Low to moderate	3.2 †	1.3	4.8	4.2
Elevated	--	--	--	--
Number of unweighted sample cases	984	1,096	521	306

Table 14-11 (continued)

	Lifetime only		No victimizations	
	ACASI ^a	CATI ^b	ACASI ^a	CATI ^b
Interviewer assessment of respondent emotional state				
Neutral	98.9 %	99.8 %	98.6 %†	100.0 %
Low to moderate	1.1 !	0.2 !	1.4	--
Elevated	--	--	--	--
Number of unweighted sample cases	181	459	282	331

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

*Comparison group.

-- Less than 0.05%

! Interpret with caution. Coefficient of variation is greater than 50%.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

14.2.2 Signs of Distress and Cooperation

Interviewers received extensive training on how to identify signs of distress by listening for verbal cues (e.g., sounds of crying, sniffling, etc.) and observing non-verbal cues (e.g., low pauses, teary eyes, etc.). As a measure of precaution, all respondents were provided a list of resources in the event they needed to seek assistance, and interviewers had detailed procedures outlining how to handle elevated distress (see Chapter 4 for more details). Less than 2.5 percent of all respondents exhibited any signs of distress regardless of sample type, but there were notable mode differences.

Almost all respondents were rated by interviewers as cooperative, regardless of mode or sample type (tables 14-12, 14-13). Very few respondents exhibited either verbal (1.6% ACASI; 2.4% CATI) or nonverbal distress (0.7% ACASI; 1.9% CATI) at any point during the interview. CATI interviewers were significantly more likely to note that respondents displayed signs of verbal distress than ACASI interviewers, as one might expect from the results described in section 14.1 above. There were differences in the amount of distress displayed by those reporting a 12-month victimization. Verbal distress was more likely to be detected among CATI respondents, as one might expect given this mode of administration is primarily based on oral communication. Among the 12-month victims, interviewers reported 9.3 percent of CATI respondents displaying verbal distress compared to the overall average for CATI of 2.4 percent. Interviewers noted that both the CATI and ACASI respondents who had reported a 12-month victimization displayed higher levels of non-verbal distress.

14.2.3 Someone Else Present During the Interview

One of the issues administering the interview in a household setting is being able to maintain the privacy of the interview. It can be difficult to control the setting of an interview that is conducted in a private residence. Space within the housing unit can constrain the extent to which the interview can be done in private, as well as keeping the respondent-interviewer interactions out of the earshot of others in the household. Several precautions were taken to facilitate this. One was the graduated consent process, which did not reveal the topic of the survey until talking directly to the respondent. For the ACASI, the informed consent, which revealed the topic of the survey, was done as a self-administered module. Also for the ACASI, the privacy of the questions was maintained as long as no one was able to view the computer screen. For the CATI, the interviewer informed the respondent about the importance of keeping the interview private. They were told they could hang up at any time they felt their privacy was threatened. Also, the response categories to the

questions were converted to “yes/no” questions. This made it difficult for anyone to understand what was being asked by overhearing the respondent’s answers.

Interviewers received extensive mode-specific training on privacy in order to protect the sensitive nature of the survey. For ACASI, the interviewers were trained to set up the interview in a private location. They were instructed to monitor, to the extent possible, that no one was in the room or at least out of earshot. They also ensured that no one was in the line of sight of the computer screen and were trained to stop the survey when someone else was looking at the screen to ensure the respondent’s confidentiality. Home office staff reviewed the initial 10 cases of each interviewer and provided corrective feedback when it appeared more privacy was needed. Afterward, cases were randomly monitored throughout data collection to ensure data integrity and confidentiality. The debriefing items related to these privacy conditions were monitored. If one was found to be an issue, the interviewer notes were reviewed and the interviewer was contacted to understand the context of the interview, as well as making sure the interviewer understood the conditions that needed to be enforced.

Table 14-12. Interviewer ratings of cooperativeness and distress, by type of victimization and mode of interview for females ages 18-49 in the general population, 2014-2015

	Total		Any past 12 months	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
How cooperative was the respondent?				
Very	90.9 %†	98.7 %	94.3 %	96.7 %
Fairly	8.6 †	1.3	5.1	3.3
Not very	0.3	--	0.6 !	--
Hostile	0.2 !	--	--	--
Did the respondent show any signs of nonverbal distress?				
Yes	1.6 %	2.4 %	6.4 %	10.4 %
Did the respondent show any signs of verbal distress?				
Yes	0.7 %†	1.9 %	1.6 %†	9.3 %
Number of weighted sample cases	11,257,283	10,524,547	1,474,828	934,526

Table 14-12 (continued)

	Lifetime only		No victimization	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
How cooperative was the respondent?				
Very	94.6 %†	98.9 %	89.1 %†	98.9 %
Fairly	5.4 †	1.1	10.3 †	1.1
Not very	--	--	0.3	--
Hostile	--	--	0.3 !	--
Did the respondent show any signs of nonverbal distress?				
Yes	0.9 %†	2.8 %	0.9 %	0.7 %
Did the respondent show any signs of verbal distress?				
Yes	0.6 %†	1.7 %	0.5 %	0.7 %
Number of weighted sample cases	2,263,988	4,437,087	7,518,467	5,152,934

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

-- Less than 0.05%

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

Table 14-13. Interviewer ratings of cooperativeness and distress, by type of victimization and mode of interview for females ages 18-29 in the volunteer sample, 2014-2015

	Total		Any past 12 months	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
How cooperative was the respondent?				
Very	97.1 %†	98.8 %	97.3 %	97.4 %
Fairly	2.8 †	1.1	2.7	2.3
Not very	0.1 !	0.1 !	--	0.3 !
Hostile	--	--	--	--
Did the respondent show any signs of nonverbal distress?				
Yes	2.1 %	2.4 %	3.3 %†	6.5 %
Did the respondent show any signs of verbal distress?				
Yes	0.5 %	1.2 %	0.8 %†	3.6 %
Number of weighted sample cases	972	1,096	515	306

Table 14-13 (continued)

	Lifetime only		No victimization	
	ACASI ^a	CATI ^{b*}	ACASI ^a	CATI ^{b*}
How cooperative was the respondent?				
Very	96.7 %	99.3 %	97.1 %†	99.4 %
Fairly	3.3	0.7 !	2.5	0.6 !
Not very	--	--	0.4 !	--
Hostile	--	--	--	--
Did the respondent show any signs of nonverbal distress?				
Yes	0.6 %!	1.3 %	0.7 %!	--
Did the respondent show any signs of verbal distress?				
Yes	0.6 %!	0.4 %!	-- %	--
Number of unweighted sample cases	180	459	277	331

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

-- Less than 0.05%

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

† Significant difference from CATI at the 95% confidence level.

^aAudio computer-assisted self-interview.

^bComputer-assisted telephone interview.

CATI interviewers were asked if they thought someone had been listening in on part of the interview. This could have occurred when CATI interviewers overhead children or adults in the background or someone verbally interacted with the respondent while she was on the phone taking the survey. For the general population sample, CATI interviewers reported that 4.9 percent of respondents may have had someone listening in on at least part of the survey. A slightly lower percentage were suspected to be listening for those who reported a lifetime victimization compared to those who did not report any victimization (3.1% vs. 6.5%) (table 14-14). However, this pattern did not occur for those reporting a 12-month victimization, where about as many people were listening as those not reporting a victimization. One would expect that reporting a 12-month victimization is much more sensitive than reporting a lifetime victimization.

For about 36 percent of the interviews someone was suspected to have listened for at least 50 percent of the time ($16.1\% + 19.9\% = 36.0\%$). Note this figure is the percent of those interviews that someone was suspected to have listened. Once accounting for the percent where anyone listened at all (4.9%), this translates to about 1.8 percent of the CATI respondents ($.36 \times .049 = 1.8$).

For the volunteer sample, 2.6 percent of interviewers felt someone may have been listening. This was not related to whether someone reported a victimization (table 14-15).

Approximately one-third of all ACASI interviewers documented instances of someone else in or entering the room during a portion of the survey. For the general population sample (table 14-16), the most common type of person in the room was a child (59.8%). This was followed by another adult (25.4%) and the respondent's spouse or partner (25.3%). When someone was in the room for at least 5 minutes, 84.5 percent of those in the room were there while the respondent was responding to the ACASI questions. In 7.7 percent of the interviews where someone was in the room for at least 5 minutes, interviewers reported that someone looked over the respondent's shoulder at some point during the interview (1.6% of all respondents).

When others looked over the respondent's shoulder, the interviewer immediately closed the laptop and explained the requirement for privacy. When others were within earshot of the interviewer and respondent, the interviewer requested that they move out of hearing distance. In all cases when an interview could not be completed due to privacy concerns, interviewers were trained to end the interview. Field supervisors monitored such occurrences and provided refresher training to ensure the interviewer intervened appropriately to maintain the privacy of the respondent.

There was no clear relationship between victimization status and any of these conditions. The only significant relationship is that those reporting a 12-month victimization were less likely to

have a child in the room than those that did not report any type of victimization (48.2% vs. 60.5%). Other relationships and the behavior of the other person were not associated with victimization.

For the volunteer sample (table 14-17), the patterns were very similar. About one-third of the interviews were done with someone in the room for any part of the interview. When someone else was in the room, it was mostly another adult. Given this sample was composed of young women, this is not too surprising as many were not married and did not have any children. Someone was in the room less often for those that reported a lifetime victimization in comparison to those with no victimizations (20.6% vs. 29.2%). However, having someone else in the room was most common for those that reported a 12-month victimization. Similarly, there was a very mixed relationship between presence, who was present and how long they were present with victimization status. For example, the proportion of interviews done when someone was in the room for at least 5 minutes was highest for those who reported a 12-month victimization (74.7% vs. 60.5%). But this relationship was reversed for some of the specific portions of the interview.

14.2.4 ACASI Survey Setting and Verbal Interaction with Interviewer

As previously mentioned, conducting the ACASI survey in a private area was important to ensure confidentiality. The ability to implement this in a residential setting, which is constrained by the physical layout and the cooperation of others in the household can be challenging. Very few respondents—only 4.9 percent of the general population (table 14-18) and only 3.4 percent of the volunteer sample (table 14-19)—expressed resistance to conducting the interview in a private setting. In the majority of instances, it was the respondent who resisted the private setting (61.0% GP, 93.9% VO), rather than a spouse or other household member. In about half of all cases, the respondent did not wear the headphones that were available to use during the interview and read the survey herself. Over half of the general population and approximately one-third of volunteer participants took the survey in a room without any doors.

For the general population sample, resistance to the private setting was associated with lower rates of any past 12-month and lifetime victimization compared to no victimization (1.3% and 3.4% vs. 6.1%). None of these factors were associated with victimization in the volunteer sample.

Table 14-14. Interviewer ratings of the presence of others during the CATI interview, by type of victimization for females ages 18-49 in the general population, 2014-2015

	Total	Any past 12 month	Lifetime only	No victimization*
Did it seem like someone may have been listening in on any part of the interview?				
Yes	4.9 %	5.1 %	3.1 %†	6.5 %
How much of the time did it seem like someone was listening to the interview? ^a				
100% of the time	16.1 %	9.5 %!	9.4 %!	19.8 %
50-99% of the time	19.9	37.4	30.9	12.9
25-49% of the time	20.9	22.1 !	10.8 !	24.8
Less than 25% of the time	43.1	31.0	48.9	42.4
Total number of weighted interviewer observations	10,524,547	934,526	4,437,087	5,152,934

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from CATI at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBase is those who said it seemed like someone was listening to the interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-15. Interviewer ratings of the presence of others during the CATI interview, by type of victimization for females ages 18 or older in the volunteer sample, 2014–2015

	Total	Any past 12 month	Lifetime only	No victimization*
Did it seem like someone may have been listening in on any part of the interview?				
Yes	2.6 %	1.6 %	2.0 %	4.2 %
How much of the time did it seem like someone was listening to the interview? ^a				
100% of the time	28.6 %	s %	s %	42.9 %
50-99% of the time	21.4	s	s	7.1 !
25-49% of the time	28.6	s	s	28.6
Less than 25% of the time	21.4	s	s	21.4 !
Total number of unweighted interviewer observations	1,096	306	459	331

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

s Data suppressed for disclosure reasons.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBase is those who said it seemed like someone was listening to the interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-16. Interviewer ratings of the presence of others during ACASI, by type of victimization for females ages 18-49 in the general population, 2014-2015

	Total	Any past 12 month	Lifetime only	No victimization*
Was anyone in the room during any part of the interview?				
Yes	35.1 %	34.6 %	35.5 %	35.1 %
Who else was in the room during the interview? ^a				
The respondent's mother or father (or both)	14.8 %	19.8 %	14.0 %	14.1 %
The respondent's spouse or partner	25.3	19.9	25.7	26.2
Some other adult	25.4	34.5	22.8	24.5
A child/children	59.8	48.2 †	65.1	60.5
When someone was in the room, was this because.... ^a				
The person walked through the area	24.0 %	27.2 %	26.1 %	22.7 %
The person was sitting or standing in the room for less than 5 minutes	31.1	28.6	30.4	31.7
The person was sitting or standing in the room for at least 5 minutes	60.8	64.4	58.8	60.7
During which portion of the interview was someone else in the room? ^b				
While asking the CAPI questions (for at least 3 questions)	85.8 %	85.0 %	79.4 %	87.8 %
While administering the event history calendar (for all or most of the time)	75.4	85.0	72.9	74.5
While the respondent was completing the ACASI (for at least 5 minutes)	84.5	86.1	89.8	82.7
While debriefing and collecting information for the incentive (for at least 2 questions)	68.2	77.4	71.3	65.8
At any point during the ACASI interview, did anyone seem to be looking over the respondent's shoulders at the questions on the laptop? ^c				
Yes	7.7 %	8.0 %!	7.2 %	7.8 %
No	92.1	92.0	92.8	91.9
Not sure	0.2 !	--	--	0.4 !
Total number of weighted interviewer observations	1,044,720	125,562	221,382	697,776

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from no victimization at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

^aBase is those who experienced anyone else in the room at any point during the interview.

^bBase is those who experienced anyone else in the room for at least 5 minutes.

^cBase is those who experienced anyone else in the room for at least 5 minutes during the ACASI interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-17. Interviewer ratings of the presence of others during ACASI, by type of victimization for females ages 18-29 in the volunteer sample, 2014-2015

	Total	Any past 12 month	Lifetime only	No victimization*
Was anyone in the room during any part of the interview?				
Yes	30.5 %	34.6 %	20.6 %†	29.2 %
Who else was in the room during the interview? ^a				
The respondent's mother or father (or both)	8.4 %	6.2 %	8.1 %!	13.6 %
The respondent's spouse or partner	6.4	5.1	8.1 !	8.6
Some other adult	70.6	76.4 †	62.2	61.7
A child/children	29.1	26.4	37.8	30.9
When someone was in the room, was this because? ^a				
The person walked through the area	13.9 %	14.0 %	5.4 %†!	17.3 %
The person was sitting or standing in the room for less than 5 minutes	30.1	27.0	40.5	32.1
The person was sitting or standing in the room for at least 5 minutes	69.3	74.7 †	62.2	60.5
During which portion of the interview was someone else in the room? ^b				
While asking the CAPI questions (for at least 3 questions)	77.5 %	70.8 %†	73.3 %	92.3 %
While administering the event history calendar (for all or most of the time)	71.9	64.6 †	73.3	84.6
While the respondent was completing the ACASI (for at least 5 minutes)	86.5	83.3	86.7	92.3
While debriefing and collecting information for the incentive (for at least 2 questions)	73.0	68.8	80.0	76.9
At any point during the ACASI interview, did anyone seem to be looking over the respondent's shoulders at the questions on the laptop? ^c				
Yes	2.6 %!	5.0 %!	-- %	-- %
No	97.4	95.0 %	100.0	100.0
Not sure	--	--	--	--
Total number of unweighted interviewer observations	77	40	13	24

Note: Estimates are based on unweighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from no victimization at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

^aBase is those who experienced anyone else in the room at any point during the interview.

^bBase is those who experienced anyone else in the room for at least 5 minutes.

^cBase is those who experienced anyone else in the room for at least 5 minutes during the ACASI interview.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-18. Interviewer ratings of what type of setting the ACASI interview was conducted in, by type of victimization for females ages 18-49 in the general population, 2014-2015

	Total	Any past 12 month	Lifetime only	No victimization*
In what type of setting was the interview conducted?				
A room with all doors closed	12.4 %	15.3 %	10.9 %	12.2 %
A room with doors, but at least one door was not closed	15.7	16.7	15.2	15.6
A room without any doors	53.4	49.5	54.5	53.8
Other setting	18.6	18.5	19.3	18.4
Was there any resistance to the request to conduct the interview in a private setting?				
Yes	4.9 %	1.3 %†	3.4 %†	6.1 %
Who resisted the request?^a				
The respondent	61.0 %	s %	66.8 %	60.0 %
The respondent's mother or father (or both)	16.4	s	12.9 !	17.1
The respondent's spouse or partner	10.3	s	6.8 !	10.8
Some other adult	11.3	s	--	13.6
A child/children	16.3	s	13.5 !	16.3
Did the respondent wear the headphones during the entire ACASI instrument, for part of it, or for none of it?				
All of interview	46.4 %	45.5 %	44.8 %	47.1 %
Part of interview	4.8	6.0	5.5	4.3
Did not wear headphones at all	48.8	48.5	49.7	48.6
Total number of weighted interviewer observations	11,257,283	1,474,828	2,263,988	7,518,467

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from no victimization at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%

s Data suppressed for disclosure reasons.

^aBase is those who experienced any resistance to the request to conduct the interview in a private setting.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-19. Interviewer ratings of what type of setting the ACASI interview was conducted in, by type of victimization for females ages 18-29 in the volunteer sample, 2014-2015

	Total	Any past 12 month	Lifetime only	No victimization*
In what type of setting was the interview conducted?				
A room with all doors closed	26.6 %	26.2 %	28.9 %	26.0 %
A room with doors, but at least one door was not closed	14.8	14.8	12.2	16.6
A room without any doors	36.0	36.7	32.2	37.2
Other setting	22.5	22.3	26.7	20.2
Was there any resistance to the request to conduct the interview in a private setting?				
Yes	3.4 %	3.5 %	1.7 %!	4.3 %
Who resisted the request?^a				
The respondent	93.9 %	100.0 %	s %	91.7 %
The respondent's mother or father (or both)	3.0 !	--	s	--
The respondent's spouse or partner	3.0 !	--	s	8.3 !
Some other adult	--	--	s	--
A child/children	3.0 !	--	s	--
Did the respondent wear the headphones during the entire ACASI instrument, for part of it, or for none of it?				
All of interview	49.0 %	51.1 %	46.1 %	46.9 %
Part of interview	4.6	5.0	3.9	4.3
Did not wear headphones at all	46.4	43.9	50.0	48.7
Total number of unweighted interviewer observations	972	515	180	277

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

! Interpret with caution. Coefficient of variation is greater than 50%.

-- Less than 0.05%.

s Data suppressed for disclosure reasons.

^aBase is those who experienced any resistance to the request to conduct the interview in a private setting.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Interviewers were required to document any verbal interaction with the respondent during the ACASI portion of survey administration. Interviewers noted if the respondents asked questions and the focus of the questions. Most respondents did not interact with the interviewer during the course of the interview; only 20.7 percent of the general population and 13.8 percent of volunteer samples did. Very few respondents seemed to refer to the event history calendar during the interview.

There were very few significant differences of interacting with the interviewer and victimization. For the general population, those reporting a lifetime victimization were more likely to ask “other” questions when compared to those not reporting a victimization (table 14-20). This rate was also elevated for those reporting a 12-month victimization, although this was not statistically different from those that did not report a victimization. For the volunteer sample (table 14-21), asking the interviewer any question was associated with past 12-month victimization rates when compared to reporting no victimization (17.7% vs. 11.2%).

Table 14-20. Interviewer ratings of respondent ACASI and event history calendar usage, by type of victimization for general population ages 18-49 (weighted)

	Total	Any past 12 month	Lifetime only	No victimization*
Did the respondent ask you any questions at any time while completing the ACASI?				
Yes	20.7 %	21.4 %	19.1 %	21.1 %
What types of questions did the respondent ask? ^a				
Computer related	44.2 %	45.6 %	40.6 %	44.9 %
Questionnaire related	55.3	60.2	56.1	54.1
Participation related	15.0	14.3	11.1	16.2
Other questions	14.4	19.3	27.2 †	10.0
Did the respondent ever look at the event history calendar during the ACASI interview?				
Yes	4.0 %	6.9 %	3.2 %	3.7 %
No	78.4	75.5	79.4	78.7
Did not notice	17.6	17.5	17.4	17.6
Total number of weighted interviewer observations	11,247,562	1,474,828	2,263,988	7,508,746

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from no victimization at the 95% confidence level.

^aBase is those who asked any questions at any time while completing the ACASI.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

Table 14-21. Interviewer ratings of respondent ACASI and event history calendar usage during ACASI, by type of victimization for volunteer sample (unweighted)

	Total	Any past 12 month	Lifetime only	No victimization*
Did the respondent ask you any questions at any time while completing the ACASI?				
Yes	13.8 %	17.7 %†	6.7 %	11.2 %
What types of questions did the respondent ask? ^a				
Computer related	32.8 %	29.7 %	41.7 %	38.7 %
Questionnaire related	52.2	57.1	41.7	41.9
Participation related	14.2	13.2	25.0 !	12.9
Other questions	17.9	15.4	33.3	19.4
Did the respondent ever look at the event history calendar during the ACASI interview?				
Yes	10.8 %	12.6 %	8.3 %	9.0 %
No	74.7	73.0	75.0	77.6
Did not notice	14.5	14.4	16.7	13.4
Total number of unweighted interviewer observations	972	515	180	277

Note: Estimates are based on weighted data. See Appendix A for standard errors.

* Comparison group.

† Significant difference from no victimization at the 95% confidence level.

! Interpret with caution. Coefficient of variation is greater than 50%.

^aBase is those who asked any questions at any time while completing the ACASI.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

14.2.5 Discussion

The purpose of the interviewer observations was to document the extent to which the interviewers observed signs of distress and the extent to which they were able to adhere to the study's privacy and confidentiality conditions. Overall, interviewers noted that very few respondents exhibited signs of distress at any point during the interview – around 2.5 percent regardless of sample type. Of the distress that was noted, it was found to be low level and did not lead to significant breaks in the interview. None of the interviews resulted in a high level of distress, which would have required stopping the interview. More signs of distress were observed for the CATI interviews in comparison to the ACASI. Also, more distress was noted among those that reported a victimization. All of these results are consistent with those reported in section 14.1 above based on self-reports of distress by the respondent.

The survey was successful in implementing the privacy and confidentiality conditions for the interview. A relatively small percentage of the CATI interviews were done where the interviewer suspected that someone was listening in on any part of the interview. For ACASI, about one-third of the interviews were done where someone was in the room at any time during the interview. A significant percentage of these were done when a child was in the room. The interviewers were successful maintaining the privacy of the ACASI questions. A very small percentage of the interviews occurred where someone was looking over the shoulder of the respondent at any time during the interview (1.8%). These were monitored by home office staff to make sure the interviewer intervened to stop the situation.

There was mixed evidence that reporting a victimization was related to the interview setting. CATI interviewers noted fewer instances of someone listening in on the interview among respondents with a lifetime victimization than among those not reporting any victimization. However, this relationship did not hold for those reporting a 12-month victimization, the most sensitive type of report on the survey. For the ACASI, the presence of anyone being present during the interview was not related to reporting either a 12-month or lifetime victimization. In some cases, the presence of someone else was positively related to victimization. For example, for the volunteer sample, those who reported a 12-month victimization were more likely to have a person standing in the room for at least 5 minutes. It may be that the mode of communication (oral with an interviewer vs. self-administered), the environmental setting (is someone present), and who is present all interact in ways that affect the comfort level of the respondent.

While instructive, the results related to the relationship between the environmental factors and victimization are very preliminary. Many of the situational factors are driven by who the respondents are and their living situation, both of which are related to victimization risk. For example, it was observed that among those reporting a 12-month victimization, a lower percentage of interviews were conducted when a child was in the room. But as noted in a previous chapter, being married is one of the strongest correlates of risk of rape and sexual assault. A multivariate analysis, controlling for the characteristics of respondents, is needed to fully assess the effects the interviewing environment may have on reporting victimization, especially in those situations where extra precautions are taken to preserve privacy and confidentiality.

15. Conclusions and Recommendations

This project had two basic goals. One was to make recommendations to improve the measurement of rape and sexual assault within the NCVS program. The second goal was to contribute to the knowledge of the best methods to collect data on rape and sexual assault. These goals were addressed by developing two designs. One design used the best practices associated with collection of data on rape and sexual assault for the NCVS program. This included the use of a self-administered questionnaire (ACASI) to maximize privacy and an in-person contact to maximize response rates. The second design was similar to those used by public health researchers, which have used telephone interviews based on a random digit dial (RDD) sample frame. The purpose of this chapter is to summarize the results of key portions of the analyses and make recommendations for the measurement of rape and sexual assault for the NCVS program.

15.1 Summary of Results

This project compared two different designs and assessed their relative quality and cost. As discussed in Chapter 1, prior research has pointed to a number of design features related to collection of RSA data that are thought to affect data quality. These include the mode of the interview and privacy of the interview, the response rate and coverage of the survey, the measurement of the key concepts and the need to minimize emotional trauma of the respondent.

15.1.1 Mode and Privacy of the Interview

The project was designed to assess whether a self-administered survey produced significantly higher rates of victimization than an interviewer-administered survey. Prior research has found that a self-administered instrument produces higher disclosure of sensitive behavior than one administered by an interviewer (e.g., Kruttschnitt, 2014; Tourangeau & Yan, 2007). For this project, it was expected that the increased anonymity of the ACASI would produce higher estimates than CATI. For the general population sample, the ACASI mode in this study did produce nominally higher estimates of rape and sexual assault within the last 12 months than the CATI mode. The incidence estimate for rape for the ACASI design was 51.0 per 1,000 women age 18-49 compared to 43.1 for the interviews conducted using the CATI (see Chapter 7). The estimates of prevalence, expressed as the percentage of women age 18 to 49 who had experienced rape or sexual assault in the prior 12-

month period, were 5.9 for ACASI and 5.3 for CATI. However, neither of these differences were statistically significant. Furthermore, there is evidence that the nominal difference is related to differential coverage of the ABS and RDD frames (see section 15.1.2 below; Chapter 10), as well as higher external telescoping in the ACASI (Chapter 9). There was some indication that mode has a greater effect among individuals who experience more incidents. For the sample of volunteers who had the highest risk of victimization, there was a significant effect of mode in the expected direction (i.e., higher rates for ACASI).

With a few isolated exceptions, the nature and characteristics of the reported incidents also did not differ by mode (see Chapter 8). For example, the percentage of incidents in which the offender is known to the victim is the same across modes. The extent of injuries that occur, where they occur, and whether the police are notified about the incident, were also the same across modes.

The vignette analysis did not find a significant difference between the two modes. ACASI and CATI respondents seem to define the different vignettes in similar ways. The two modes also performed similarly when looking at how situational factors affected the ways that respondents answered several of the victimization questions. For example, they displayed similar effects when looking at how the relationship between the victim and offender affected answers to the question on completed rape. There were no differences between modes in the estimated reliability of the sexual victimization screener items, as measured by the reinterview. Similarly, mode did not affect whether or not respondents gave a consistent response to several of the items on the detailed incident form.

The absence of a significant mode effect for this study may reflect the precautions taken to ensure privacy and confidentiality in both modes of interviewing. Only one person in the household was sampled for the survey and a graduated consent procedure was used so no one but the respondent knew what was asked on the survey. For the CATI survey, respondents were asked to make sure no one was listening and to hang up if they were apprehensive about anyone knowing what was on the survey. The CATI survey questions primarily had “yes/no” response categories so anyone within earshot could not tell what was being asked. The lack of a significant difference between the modes suggests that privacy concerns related to asking about RSA are related to disclosure to others in the household. Privacy concerns related to disclosing to the interviewer did not seem to affect responses.

One notable difference between the modes is the amount of time that was taken to complete the interview. The ACASI respondents took significantly less time when filling out the detailed incident form when compared to the CATI respondents. On average, it took CATI respondents

about 8 minutes longer, compared to ACASI respondents, when completing an interview with one detailed incident form (DIF). For those having to fill out a two or more DIFs, the CATI interview took, on average, about 20 minutes longer. This increased time was also reflected in the number of respondents who completed all of the DIFs that were requested. For the ACASI, well over 90 percent filled out all of the forms requested. This percentage declines to as low as 60 percent for CATI respondents who were asked to fill out three incident forms.

The longer administration time on the CATI is due to the need to read out all of the response categories for each question. Respondents were asked to respond yes or no for each question or were asked to reply in the form of a categorical number rather than repeating the substantive response out loud to the interviewer. This was done to prevent anyone who could overhear the respondent from knowing about the content of the survey. The self-administration mode lends itself better to collecting details about the incident. Respondents can be presented with lists of the possible options without overly burdening them by reading out each response alternative. The visual presentation of response options, as can be done for self-administration, also lends itself better to the task of asking about the circumstances related to incidents. This is not possible over the telephone, even if open-ended questions are asked without reading the response categories.

15.1.2 Response and Coverage Rates

One question related to data quality is whether the estimates from one or both of the surveys had significant non-response bias. Non-response bias is one source of error that has been discussed when comparing the NCVS, which has a relatively high response rate, to many of the RDD surveys, which generally have lower response rates. To examine this question, the RSA Pilot Test compared the effects of non-response on the ACASI and CATI estimates. The ACASI was based on in-person contact and achieved a response rate of approximately 40 percent. The sample frame for the CATI survey was RDD and achieved a response rate of 18 percent (Chapter 3). Analyses did not find evidence of significant non-response bias in either survey. Using census characteristics for the areas where the respondents were located showed no indications of significant under- or over-representation of demographic or economic groups for the surveys conducted with either mode of interview (Chapter 10). A level of effort analysis did not find a correlation between the number of contacts needed to complete the survey and reporting victimization. These conclusions are consistent with several recent campus climate surveys (Cantor et al., 2016; Krebs et al., 2016), as well as early analysis of the NISVS (Peytchev et al., 2009). These surveys also did not

find evidence of significant non-response bias for outcomes related to rape or sexual assault among surveys with response rates between 19 percent and 40 percent.

As with most evaluations of non-response, the above analysis cannot definitively conclude that estimates from either the ACASI or CATI data collections are free of bias due to nonresponse. The evidence from the level-of-effort analyses, cited above, are based on the assumption that the late responders to the RSA Pilot Test surveys are representative of the non-responders. In particular, the analysis assumes that the 60 percent of the sample that did not respond to the ACASI and the 82 percent that did not respond to the CATI survey have similar rates of rape and sexual assault as the late responders to the survey. There was no way to test this assumption with the data from this study. Several studies of other types of outcomes have shown that this assumption does not always hold (Lin & Shaeffer, 1995). Future research, perhaps with surveys that have a higher response rate (e.g., the NCVS) or have data on non-responders, would shed more light on the role non-response may play on surveys of rape and sexual assault.

The CATI survey was based on an RDD sample frame, which did not allow precise targeting of sampled cases within the sampled CBSAs. The frame is based on area codes and there are a significant number of individuals with mobile phones who have area codes from outside the CBSA. For example, approximately 10 percent of the ACASI respondents who owned a mobile phone did not have an area code that was included on the RDD frame. Analysis of this group of respondents found them to have significantly higher rates of rape than the general population. This suggests that when targeting a particular CBSA, an RDD sample frame may undercover victims of rape. Those residents who are missed because their area codes are outside the sample frame were more likely to report a rape than those who fell into the sample frame.

15.1.3 Missing Data

With a few exceptions, item-missing data was generally low (below 5 percent) for both modes of interviewing. For example, the missing rate for the victimization screener items was less than 1 percent. For most of the items on the detailed incident form there was similarly low rates of missing data. For the ACASI, 87 percent of the questions had less than 5 percent missing. The percentage was slightly higher for the CATI (94%). One set of items with more missing data were the questions that asked for detail on the “Other (specify)” response categories. Another set were the questions asking about unwanted behavior and type of physical force, which had missing rates between 6 percent and 14 percent. The most missing data for the DIF were requests for the

narrative collected at the end of incident form, with about 30 percent of incidents missing a narrative.

The willingness of respondents with multiple victimizations to complete all of the assigned detailed incident forms decreased as the number of incidents increased (Chapter 9). For example, among those reporting a single incident, 99 percent and 98 percent, respectively, of the ACASI and CATI respondents completed a detailed incident form. This rate drops to 98 percent and 93 percent for ACASI respondents who had at 2 and 3+ eligible incidents, respectively. This drops much more dramatically for the CATI, with 83 and 65 percent filling out all eligible forms, respectively. The higher rate of missing DIFs for the CATI is attributed to the additional time it took to administer the DIF for this mode. The additional time was due to the requirement that all questions were in a “yes/no” format.

15.1.4 Measurement Error

Measurement error refers to the extent that a particular response corresponds to the true value (Groves et al., 2004). This project did not have an independent true value to compare to the survey responses to directly measure this type of error. However, it did collect a number of indicators that provided information on measurement error related to the survey instrument. Data quality was evaluated by various methods and included—information from the detailed incident form, approximately 1,000 re-interviews, vignettes administered at the end of the interview, debriefing information provided by respondents, and interviewer observations.

False Positives

The use of BSQs, multiple screening items, and increasing the privacy of the interview are intended to reduce false negatives (underreporting). As discussed in Chapter 1, one possible problem with this methodology is that this may come at the cost of increasing false positives (overreporting). A false positive is an incident reported on the survey that is outside of the scope of the survey (e.g., outside the reference period; does not meet criteria of a rape or sexual assault). By asking multiple questions about related behavior, respondents may feel pressure to answer in the affirmative. For example, the respondent may report an incident that occurred outside the reference period or may include an incident even though they are unsure whether it meets the criteria related to the question (e.g., does not involve clear-cut use of force or inability to consent).

One purpose of a two-stage design is to evaluate whether specific incidents meet the operational criteria used to define a rape or sexual assault. For each incident reported on the RSA Pilot Test, a series of questions were asked about the use of physical force and the respondent's ability to provide consent. These are two essential elements for defining an incident as a crime. With respect to physical force, about 80 percent of incidents that were classified as a completed rape included some type of physical force, such as being held or pinned, the use of a weapon or threat with a weapon, a physical attack or threat of physical attack without a weapon,⁵¹ a physical attack or threat to attack someone else, or being blocked or otherwise prevented from leaving (e.g., locked in, handcuffed). All of these tactics clearly qualify as force in the context of a completed rape.

On the RSA Pilot Test, about 20 percent of completed rape incidents were those in which no physical force was used, but the victim was classified as being unable to consent.⁵² Incidents involving inability to consent are sometimes difficult to prove in a court proceeding and pose unique problems with respect to measuring on a survey. The vignette analysis found that respondents tended to have an overly broad interpretation of "inability to consent." When presented with vignettes in which victims were drinking, there were significant number of respondents who reported she was not able to consent regardless of the level of intoxication. The data from the DIF is not entirely consistent with this. Most victims of incidents that met the criteria for inability to consent also reported signs of being intoxicated to the point they had trouble making decisions and/or provided an indication that another individual would recognize the victim was not in condition to consent.⁵³ Slightly more than half of the incidents classified as "rape due to inability to consent" involved victims who were unconscious for at least part of the incident. In slightly less than half of the incidents, the victim was conscious but said she was unable to consent because of alcohol or drugs. Approximately 90 percent of these victims also reported one or more additional signs of inability to consent, such as not being able to communicate with others, not being able to walk by herself, being less able to physically resist, and having the perpetrator continue to give her alcohol or drugs after she was clearly drunk or high (Chapter 8).

While there is still some ambiguity related to a self-reported measure of ability to consent, the above data does not indicate a significant problem with false positives. The vast majority of

⁵¹This includes a 'yes' response to a direct question of whether physical force was used, as well as the respondent saying she was hit, punched, bitten, choked, slapped, kicked, had her mouth covered, grabbed, pushed, pulled or groped.

⁵²There were a number of incidents where both force and inability to consent occurred. The study counted these incidents as forcible rape.

⁵³This is a key element in prosecuting alcohol/drug-related rape cases.

those who were classified as a completed rape or sexual assault victim showed signs of impaired judgement or physical impairments.

The RSA Pilot Test also included questions about how victims reacted during the incident. These items provide a picture of how the victim expressed non-consent during the incident. For incidents involving kissing, groping, or other type of sexual touching, if the respondent reported the perpetrator immediately stopped after she said “no,” the incident was not counted as a sexual assault, regardless of the tactic or ability to consent. These data also reveal that the vast majority of the victims of rape and sexual assault did express non-consent during the encounter. Of the incidents classified as a rape, 84 percent of victims physically resisted or tried to physically resist,⁵⁴ 92 percent said “no” or “stop,” 68 percent pleaded or argued with the perpetrator and 64 percent tried to escape or get away.

A final piece of evidence related to false positives is the review of narratives. Generally speaking, these descriptions were consistent with the information used in the classification of the incident as a completed rape or sexual assault (see Chapter 7). While some false positives were found, there were relatively few. The classification into attempted or threatened rape was more problematic. Respondents tended to have a broader definition of what constituted an attempt or threat than the legal criteria. The classification scheme used in the analysis made adjustments for some of these incidents (Chapter 7). Below we make recommendations on improving the measurement of attempts and threats with changes to the questions on the detailed incident form.

Error Related to One- and Two-Stage Estimates

Both the ACASI and CATI design used a two-stage survey. Error was assessed at both stages. The overall rate of victimization using the behavior specific questions (BSQ) (first stage) and the DIF (second stage) were very similar for both modes of interviewing. For example, the ACASI estimates for completed rape from the BSQ and the DIF were, respectively, 2.8 and 2.4 percent. However, the specific incidents that classified as a completed rape were somewhat different. For example, approximately 51 percent of the incidents classified as a completed rape from the BSQ for the ACASI was classified in this category from the DIF. Similarly, approximately 70 percent of the incidents from the DIF classified as a completed rape were classified in this category from the BSQ.

⁵⁴These refer to the results for the ACASI survey. The results for the CATI are similar.

When comparing specific incidents to the narratives provided by respondents, measurement error was apparent for both types of measures.

External Telescoping

External telescoping is when respondents remember an event that occurred outside the reference period and misdate it as occurring in the reference period. For victimization surveys, this can be a significant source of error (Cantor, 1989). It is more likely to occur for salient events, which are more likely to be remembered in the first place. The ACASI survey tried to control for this type of error by administering an event history calendar before asking about victimization. This was intended to assist the respondent by anchoring incidents around landmark events. The CATI survey used an internal bounding procedure, which first asked about lifetime events and then asking whether anything had happened with the last 12 months (Sudman et al, 1984; Loftus et al, 1990).

Examination of the dates in the reference period indicated that the internal bounding procedure was more successful at controlling external telescoping than the event history calendar. For both the general population and the volunteer sample that answered via ACASI, the percentage of rapes reported at the beginning of the reference period was significantly higher than for other types of events. This is a signature that external telescoping is occurring. For the general population sample answering via CATI, there was not a significant upturn of reports of rape at the beginning of the reference period. The data for the volunteer sample was not as clearcut. The percentage of incidents reported in the first month of the reference period was about average. However, the percentage reported in the second and third months were significantly higher.

Reliability

Data reliability was examined by conducting approximately 1,000 re-interviews, two to three weeks after the initial interview, among those who reported an unwanted sexual contact at the first interview. The screener items, when grouped into logical categories, had Kappas of 60 to 70 percent; these are considered to be “substantial” using the Koch and Landis (1977) standard. Nonetheless, this is lower than one might expect given the fairly specific behaviors that are being referenced on the screener. For example, approximately 30 percent of respondents who reported a completed rape on a screener item at the first interview changed answers to something else at the second interview. The change was distributed across the remaining types of victimizations, including not reporting any

12-month incident at all. When respondents were asked why they changed responses, the most prominent reason was re-interpretation of the questions at the second interview. The next most common reason was something being remembered differently at the second interview. Some respondents said they did not report an incident to avoid being asked the follow-up questions.

The analysis of the DIF concentrated on those incidents that could be identified as being the same based on the narrative. This provided a direct measure of whether respondents were answering the DIF items the same way for the same set of target behaviors across the two interviews. Based on this subset of incidents, the questions on unwanted behaviors had lower reliabilities than expected. Some of this was related to confusion about how to define an attempted or threatened act. There also were higher than expected inconsistencies related to respondents reporting a completed act (e.g., penetration, sexual touching). Most of the latter were respondents reporting a completed act at the first interview and no unwanted behavior at the second interview. For example, of those that reported a completed vaginal penetration at the first interview, 30 percent reported that no unwanted vaginal penetration occurred at the second interview. The other items used to classify the incident, such as “using force” and “inability to consent” had relatively high agreement rates and reliabilities.

When examining the consistency of the overall crime classification algorithm, which combined items to determine the type of crime, completed rape had the highest consistency of the major categories, with approximately 70 percent of the incidents being classified the same way at both interviews. Consistent with the discussion above, many of the changes in classification between interviews were related to respondents reporting completed penetration at one interview and no type of penetration at the second interview. Review of the narratives suggests that this type of change was due to problems with the structure of the unwanted behavior questions, as the narratives indicated penetration had occurred in most cases. Sexual assault had the lowest consistency. For completed sexual assaults, the inconsistency was related to the omission of a category for groping and grabbing in the force question. The attempted and threatened acts of sexual assault were subject to a similar problem; however, there was also some indication that change between interviews was related to ambiguity with respect to what constituted a forced attempt or threat.

All of the above analyses (one-stage vs. two-stage; re-interviews) pointed to problems with asking about attempted and threatened acts. Respondents tended to have a broader definition of these types of incidents than would be expected in the definition of a crime. In particular, some respondents reported an attempt/threat of rape or sexual assault in anticipation of what might happen or in response to a verbal threat, even though the threat may not have been credible. For

example, verbal harassment was sometimes reported as a threat to rape even though the threat was not imminent or credible. The classification of attempts and threats made some adjustments for this error by using the narrative descriptions. The recommendations below provide suggestions on how to improve the measurement of these types of incidents.

15.1.5 Minimize Emotional Trauma

There were several human subject concerns addressed on the project. The first was to implement procedures to mitigate harm to respondents who may have adverse reactions to the survey. Prior research has found that respondents do not report substantial harm from taking sensitive surveys similar to the RSA Pilot Test. Even those who do report negative feelings do not typically regret taking the survey because they also report experiencing positive emotions such as empowerment or contributing to solving the problem (DePrince & Chu, 2008; McClinton et al., 2015; Newman et al., 2006). One unique feature of the RSA Pilot Test is administration of a relatively lengthy detailed incident form, which could be perceived as more intrusive than prior surveys. Findings from the RSA Pilot Test are in line with prior research, with rates for positive and negative reactions falling within ranges reported elsewhere (Black et al., 2006; Valpied et al., 2014; Wager, 2012; Walker et al., 1997). Respondents in the RSA Pilot Test reported higher agreement with items indicating more positive reactions than negative reactions, and most respondents indicated that they did not regret taking the survey. This general pattern held true across survey modes and levels of victimization.

Similarly, interviewers noted that very few respondents exhibited signs of distress at any point during the interview – around 2.5 percent regardless of sample type. The distress that was noted was found to be low and did not lead to significant breaks in the interview. None of the interviews resulted in a high level of distress, which would have required stopping the interview.

A second concern was being able to conduct the interview in an environment that preserved the confidentiality of the interview. Both the ACASI and CATI interviews were administered to maintain this privacy. In both cases, the topic of the survey was not revealed until administering the informed consent to the selected respondent. For the in-person visit, the informed consent was incorporated on the ACASI. For the telephone interview, most questions were structured so respondents only had to answer “yes” or “no” to make it difficult for anyone else in the household to understand what the interview was about. One concern for the ACASI was avoiding situations where someone else in the household was able to hear or see the interview. Especially in small

housing units (e.g., small apartments), it may not be possible to fully prevent individuals from walking into the interviewing space or overhearing from another room. Isolating respondents in a private setting may be difficult to negotiate in a household setting. For the CATI, the interviewer had less control over who might be listening to their interactions.

The survey was successful in implementing the privacy and confidentiality conditions for the interview. A very small percentage of the CATI interviews were done where the interviewer suspected that someone was listening in on any part of the interview. For ACASI, in about one-third of the interviews, someone entered the room at some point during the interview. A significant percentage of these were done when a child was in the room. The interviewers were successful in maintaining the privacy of the ACASI questions. A very small percentage of the interviews occurred where someone was looking over the shoulder of the respondent at any time during the interview (1.8%). In those cases where this occurred, the home office staff followed up with the interviewer to make sure she took the appropriate actions (i.e., stopped the interview).

15.2 Comparisons with the NCVS

The analysis in chapters 7 and 8 compared the RSA Pilot Test estimates to the NCVS along several dimensions. In this section, the results and implications of these comparisons are discussed.

15.2.1 How do RSA Pilot Test Estimates Differ from those of NCVS?

The incidence and prevalence estimates for the RSA Pilot Test are higher than the NCVS estimates by a factor of 50. One significant difference between the two surveys is the wider scope of incidents which are cued on the RSA Pilot Test screener. The RSA Pilot Test included 14 screening items, covering a wide range of behaviors and tactics. These cues serve to both prompt recall of specific incidents, as well as to define eligible incidents. The NCVS, in contrast, includes two screening questions that target rape and sexual assault. When restricting the RSA Pilot Test estimates to include the types of incidents that are covered on the NCVS screening questions, the difference between the estimates was reduced to a factor of 18. A second difference between the surveys is the framing of eligible incidents. The NCVS is a crime survey. It includes “crime” in its title and references “crimes” in several points of the screener. In addition, the rape and sexual assault questions are asked within a series of cues that ask about other predatory crimes, such as robbery,

burglary and motor vehicle theft. The analysis assessed the effects of this context by restricting RSA Pilot Test estimates to those incidents respondents said were crimes at two different points in time (at the time of the survey; at the time of the incident). This reduced the difference between the two surveys to a factor of between 5 to 10.

The RSA study could not estimate how much other design features of the two surveys may contribute to the differences. One of these features is the extent to which privacy impacts reports. The NCVS attempts to interview everyone in the household age 12 and over. This communicates the survey questions to all household members, as well as leading to interviewing environments that involve multiple members of the household being present (Catalano, 2016). A second design feature that differs is related to the greater number of screening items that target sexual violence on the RSA Pilot Test (14 vs. 2 for the NCVS). The additional cues on related types of incidents, even if not specifically targeting a rape or sexual assault, may lead to additional recall of events. One indication of this is that a significant number of rapes (17%) and sexual assaults (25%) were reported in response to screening items that did not specifically target these particular types of incidents.

As discussed in Chapter 1, one common reason cited for the elevated rates on a survey like the RSA Pilot Test are false positives. Previous criticisms of sexual assault surveys using BSQs is that the difference with the NCVS is related to an increase in false positives. A high false positive rate may be one reason why the RSA Pilot Test estimates are higher than the NCVS. and that this is why there is such a large difference with the NCVS. As discussed in section 15.1.4, very little evidence was found that the RSA Pilot Test had a significant number of false positives for completed rapes and sexual assaults. There was some indication that attempted and threatened incidents may be subject to more false positives because respondents may use an overly broad definition of an attempt or threat than intended. The RSA Pilot Test tried to minimize this by using the narratives to identify obvious false positives, but there still may be some remaining in the final estimates.

There are several other differences between the NCVS and the RSA Pilot Test that likely contribute to the differences. One is the effect panel conditioning has on the NCVS estimates. The rates of victimization on the NCVS for respondents at later points in the panel rotation differ by significant amounts (e.g., factors of 2 to 4) when compared to the first two interviews. A second difference is the response rate. In the particular cities that were included in the RSA Pilot Test, the average response rate on the NCVS was approximately 28 percentage points higher than the RSA ACASI survey and 50 percentage points higher than the RSA CATI survey. If non-victims are more likely to participate on the NCVS relative to the RSA Pilot Test, then some of the difference between the estimates may be related to non-response bias.

15.2.2 Characteristics and Correlates of Victimization

The correlates of victimization risk were very similar for the two studies. For the RSA Pilot Test, the strongest correlates associated with both rape and sexual assault were age, marital status, and race/Hispanic origin. Females ages 18 to 24 had victimization rates for rape that were significantly higher than those just a few years older (25-29); this trend continued into the older age groups. Those who are married had significantly lower rates than those not married. With respect to race/Hispanic origin, non-Hispanic white women have the highest rates. Household income was also significantly related to rates of rape. Those in the lowest income group had the highest victimization rates. Analysis of the NCVS, both in bivariate (Planty et al., 2013) and multivariate analyses (Lauritsen, 2012), have found the same effects of age, marital status, race/Hispanic origin, and income.

A second similar finding is that women enrolled in college did not exhibit higher rates of victimization than non-college students (Chapter 8). Nationally, there is considerable concern over the high rates of sexual violence among college students, as revealed by recent campus climate surveys (White House Task Force Report, 2014). In analysis of the NCVS, Sinzit and Langton (2014) found that, after controlling for age, women age 18 to 24 who are not in college have slightly higher rates than those who are in college. The RSA Pilot Test found that college enrollment did not increase risk. Those currently in college had a similar victimization rate as those who were not in college.

The types of incidents that are captured by the NCVS and RSA Pilot Test surveys are similar along several characteristics. Both surveys enumerated incidents that infrequently involve weapons (around 8% of the time) and most involve one male offender. With respect to the relationship between victim and offender, they have the same proportion of offenders who were casual acquaintances and those known by sight only.

There are, however, several key differences in the characteristics of incidents between the two surveys. The largest difference is the percentage of incidents in which the police found out about the incident. Incidents identified on the RSA Pilot Test were three times less likely to come to the attention of the police than those on the NCVS (10% vs. 34%). One important reason incidents are reported to the police is because they are more salient, such as resulting in more consequences or injuries. This is consistent with other characteristics that differed between the two surveys. For example, NCVS respondents were more likely to report physical injuries, emotional difficulties for at

least one month, having to go to the emergency room or other hospital setting, and getting help from a victim assistance agency.

A second dimension where there was a difference was the extent the incidents on the two surveys involved intimate partners. The RSA Pilot Test had a higher percentage of friends or ex-friends and strangers as perpetrators, while the NCVS has a higher percentage of spouses and ex-spouses. This resulted in a higher percentage of incidents reported on the NCVS which involved intimate partners using the NCVS definition (spouses, ex-spouses, boyfriends and girlfriends). Related to this, more of the RSA incidents occurred at a friend's house, while more NCVS incidents occurred at the respondent's home.

15.3 Recommendations

This section presents recommendations for improving the measures of rape and sexual assault within the NCVS program. These recommendations are summarized in exhibit 15-1. The first section discusses recommendations for changes to the core survey and implementation of a more specialized survey on rape and sexual assault. The section discusses changes to the RSA Pilot Test design to improve the measurement of rape and sexual assault.

15.3.1 Recommendations for Collection of Rape and Sexual Assault Data Within the NCVS Program Design

Recommendation 1: For the ongoing NCVS, redesign the screening items that target rape and sexual assault. Expand the scope of the items to include different types of penetration and sexual contact. Expand the items on the detailed incident form to ask about the behaviors and tactics that are specific to rape and sexual assault.

The evidence from this study suggests that one of the major reasons the NCVS rates are lower than expected is the relatively narrow focus of the screening questions. It is recommended that the questions on the NCVS use specific language to describe the types of behaviors (e.g., penetration, kissing, groping) and tactics (e.g., physical force; inability to consent) that constitute the definitions of rape and sexual assault. At a minimum, two questions should be used. One question should ask about unwanted sexual contact that involved force, including attempts. A second question should ask about unwanted sexual contact that happened while the person was unable to consent. However, the RSA Pilot Test suggests that more than two questions are necessary to fully

enumerate all types of incidents that are within scope. The recommendation of two items is a minimum, keeping in mind that the NCVS is an omnibus survey that collects data on several other types of crimes as well as other topics of interest (e.g., contact with the police; school crime). If additional items can be added, then it is likely that fuller measurement can be achieved.

Exhibit 15-1. Recommendations for collecting data on rape and sexual assault within the NCVS data collection program

1. For the ongoing NCVS, redesign the screening items that target rape and sexual assault. Expand the scope of the items to include different types of penetration and sexual contact. Expand the items on the detailed incident form to ask about the behaviors and tactics that are specific to rape and sexual assault
2. For the ongoing NCVS, improve procedures to ensure interviews are conducted without any other persons present.
3. If redesign of the core NCVS does not fully measure rape and sexual assault, implement a separate survey within the NCVS program to collect these data.
4. Design the separate survey to produce incidence and prevalence estimates, as well as describing the characteristics related to rape and sexual assault. The sample sizes should support generating these estimates on a rolling, multi-year basis.
5. Investigate two methods to draw a sample for the separate survey. One is to sample from respondents to the core NCVS. The second is an independent sample.
6. Define Rape as penetration involving physical force or inability to consent. Define sexual assault as other types of kissing or other sexual contact involving physical force or inability to consent. Include attempts and threats for both rape and sexual assault.
7. The mode for the separate survey should be adapted to other features of the design, such as the sample design and methods to contact the respondent.
8. The separate survey should use behaviorally specific questions, a detailed incident form and a narrative.
9. Limit the number of detailed incident forms to three. For respondents reporting more than three incidents, assign the highest priority to incidents involving penetration, followed by those involving sexual touching. Subsample incidents if a respondent has multiple incidents involving unwanted sexual contact.
10. If the separate survey cannot use the same bounding procedure as the NCVS, use a sequence of questions that first ask about lifetime and then ask for anything occurring in the last 12 months.
11. For the separate survey, revised the RSA Pilot Test victimization screening items such that: (1) when describing penetration, the wording should communicate that very slight penetration should be counted and (2) describe non-consent in the items involving physical force as “against your will.”
12. Simplify the RSA Pilot Test questions on unwanted behavior on the detailed incident form.
13. Collect a combined measure of threats and attempts. Institute follow-up questions that operationalize how the threat or attempt was carried out for unwanted penetration.
14. Link the force questions on the detailed incident form to the type of unwanted behavior that is reported. Separate the tactic questions for penetration and other type of sexual contact.
15. Consider using multiple criteria when classifying an incident as being related to alcohol and drugs.

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

The DIF on the NCVS does not enumerate the specific behaviors or tactics that define rape and sexual assault. For example, on the NCVS a rape is operationally defined when respondents report ‘rape’ in response to the question of how they were attacked. Besides intermingling the legal terminology of rape/sexual assault with a type of attack, this approach misses descriptors of force that make up this type of tactic (e.g., pinning down). There are no questions that ask about the different behaviors, such as different forms of penetration or unwanted kissing or groping. Finally, there are no questions related to the respondent’s ability to consent. To improve the measurement, therefore, the NCVS DIF should enumerate the type of sexual contact, including the different forms of penetration, kissing, and sexual touching and whether they were completed or attempted/threatened. Similarly, the tactics should be enumerated, including the different types of physical force that may be used, any groping behaviors, and the ability to consent.

Recommendation 2: For the ongoing NCVS, improve procedures to ensure interviews are conducted without any other persons present.

A significant percentage of the interviews on the core NCVS are administered when someone else is present in the room (Catalano, 2016). For in-person interviews, there may also be individuals who are out of the room, but within earshot. This inhibits disclosure of events that respondents may not want others to know about. The results from the RSA Pilot Test demonstrated that it should be possible to increase the privacy of the interview. This can be done in several ways. One would be to change the mode of the interview on the NCVS from interviewer-administered to self-administered. Privacy can be further enhanced by training interviewers to make sure, to the extent possible, that others are not present and/or are out of vision and earshot of the interview. This should hold not just for the NCVS items on rape and sexual assault, but for all of the data collected on the survey.

About half the NCVS interviews are conducted in person and half are conducted over the telephone. Ensuring the telephone interviews are private can follow similar protocols as used on the RSA Pilot Test. Interviewers should emphasize that the respondent be in a private location without anyone else overhearing. A similar condition should be followed for in-person interviews. It is recognized that smaller housing units may not allow absolute confidentiality because it may be hard to stay out of earshot of others. If the survey is converted to self-administration, it should still be possible to maintain confidentiality. If the survey is interviewer-administered, field representatives should be trained to maximize confidentiality as much as possible, at least ensuring physical isolation. If this is not possible, field representatives should consider accepting non-response in lieu of an interview that is not private.

15.3.2 Recommended Methods and Design Features of a Separate Survey

This section makes recommendations on the collection of RSA data on a separate survey and the design features recommended for this survey.

Recommendation 3: If redesign of the core NCVS does not fully measure rape and sexual assault, implement a separate survey within the NCVS program to collect these data.

One goal of the RSA Pilot Test was to inform the decision on whether valid estimates of rape and sexual assault can be collected on the NCVS or whether RSA needs to be collected as part of a separate survey.⁵⁵ The RSA Pilot Test survey did not test a revised version of the NCVS with improved measures of rape and sexual assault. Consequently, results from this study do not directly address whether a separate survey would be needed after applying the recommendations above, i.e., expanding the scope of the NCVS screening items, making changes to the detailed incident form and improving the privacy of the interview. It will be important to test the changes to the NCVS design noted above to assess the extent to which full measurement can be achieved.

However, the RSA Pilot Test results suggest that, while it should be possible to improve the estimates on the ongoing survey, it is not likely that it can fully measure rape and sexual assault and stay within the current constraints of the NCVS design. On the RSA Pilot Test, nine screening questions were needed to fully enumerate the behaviors and tactics that define rape and sexual assault, as well as to provide respondents with the cues to assist in recall of these incidents. It may be difficult to fully enumerate all types of incidents by just modifying the core NCVS with two screening items. The analysis of the RSA Pilot Test also suggests that the crime context is an important constraint on the core NCVS.

Recommendation 4: Design the separate survey to produce incidence and prevalence estimates, as well as describing the characteristics related to rape and sexual assault. The sample sizes should support generating these estimates on a rolling, multi-year basis.

The goals of a separate survey would be to produce incidence and prevalence estimates of rape and sexual assault, characteristics of these incidents, and estimates of change over time. The precise sample size needed for this survey depends on a number of parameters (frequency of estimates, precision desired). There are approximately 220,000 individuals ages 12 and over who are

⁵⁵For purposes of discussion, it is assumed that the basic design features of the NCVS will not change. For example, it is assumed that all persons 12 and over in the household are interviewed, the omnibus nature of the survey remains the same and a significant percentage of interviews will be conducted by an interviewer.

interviewed on an annual basis for the NCVS to produce annual estimates.⁵⁶ Even with this large sample size, the precision is not high for estimates of rape and sexual assault. Breaking these data out by demographic groups or specific characteristics typically requires aggregating many years of data.

A separate survey, which implements procedures similar to the RSA Pilot Test should produce a higher rate of rape and sexual assault victimization and would not require as large a sample size as the NCVS. For example, the RSA Pilot Test estimated a prevalence rate for rape for women 18 or older to be 1.9 percent, with a relative standard error of 11 percent.⁵⁷ If the number of interviews was increased from 5,000 to 20,000, the relative standard error would be closer to 5 percent. Estimates of rape for males would be more problematic because the prevalence rates would be expected to be much lower than for females. However, it might be possible to produce estimates of sexual assault reported by males. Assuming both males and females are interviewed in equal numbers, a total of 40,000 interviews is a starting point when considering the size of the sample for the separate survey.⁵⁸

These calculations need to be refined, once parameters, such as the reference period and sample design, are defined more explicitly. The efficiency of the sample is dependent on features such as whether the survey is cross sectional or longitudinal, the extent of geographic and household clustering, the response rate, and the survey weighting.

Recommendation 5: Investigate two methods to draw a sample for the separate survey. One is to sample from respondents to the core NCVS. The second is an independent sample.

There are at least two possibilities for drawing the sample for the separate survey. One possibility is to administer the separate survey to the sample already participating on the core NCVS. One person in a household would be asked to participate in the separate survey after completing the NCVS. This is a design that is similar to what is currently done on the Crime Survey of England and Wales (CSEW). A second option is to draw an independent sample to administer the separate survey. Households or persons would be sampled using a general population frame, such as a list of housing units (ABS) or telephone numbers (RDD).

⁵⁶Note the NCVS interview uses a 6 month reference period. All else being equal, it takes twice as many NCVS interviews to cover the same calendar period as the RSA Pilot Test, which used a 12 month reference period.

⁵⁷This is the estimate from the RSA Pilot Test CATI survey once including all women 18 years and older.

⁵⁸This assumes a 12 month reference period. Sample sizes would need to be adjusted if a 6 month period was used.

One advantage of using the NCVS sample is that it efficiently targets high-risk groups at very little cost. For example, it would be useful to oversample young adults, who are at highest risk of rape and sexual assault. For a survey that draws an independent sample, oversampling would significantly add to the cost of the data collection. On the RSA Pilot Test, which only conducted in-person interviews with women in the 18 to 49 age group, the sample size was reduced by approximately 25 percent to offset the cost of screening out households. It was not efficient to conduct this screening for the RDD survey because of the marginal difference in the cost of screening households into the survey relative to the cost of completing the full interview once recruited into the sample (Chapter 3).

One disadvantage of sampling NCVS respondents is that it could impact participation in future waves of the NCVS. For example, if the separate survey is administered the first time the interviewer visits the household, it may deter participation on the NCVS at future waves. If this were true, then some consideration might be given to administering the survey at later waves (e.g., at the sixth or seventh interview). Another possible disadvantage of conducting the survey with NCVS respondents is the possibility the frame of reference from the NCVS, which is focused on crime and criminal justice, may influence how respondents to the RSA survey interpret the questions on the separate survey.

One advantage of an independent sample is that it could be more flexible in its design and procedures. For example, a sample from the NCVS would need to be integrated within the rotating panel design of the survey. Participation on the separate survey may affect response rates at the next household visit or it may influence how respondents answer the core NCVS questionnaire at the next interview.

The relative costs and response rates of the two approaches should be considered. These two parameters (cost and response rate) are directly correlated. On its face, a survey that samples from the ongoing NCVS should be less expensive than one that draws an independent sample. The costs for initial recruitment of the sample are absorbed by the ongoing NCVS. This is especially the case if groups at high risk of victimization are oversampled (see above). A recent study examining the rotating panel design of the NCVS placed the cost of completing a core NCVS interview at \$120 and \$250 by telephone and in-person, respectively (Berzofsky and Carrillo-Garcia, 2017). About half of the NCVS interviews are completed in each mode, yielding an average cost of about \$185 per complete. A separate survey that samples from the NCVS would likely be less expensive than this, since many of the interviews would be immediate follow-ups with someone completing an interview,

so no additional trips to the household will be required for these individuals. It is also possible to have the respondents complete the survey on the web, which should further reduce costs.

An independent RDD telephone survey that oversamples high risk groups (e.g., young adults) would be more expensive than this. The RSA Pilot Test costs for the RDD survey were more than \$200 per complete and this did not oversample any age groups. A second disadvantage of an independent sample using RDD is the response rate would be considerably lower than one that samples from the ongoing NCVS. The RSA Pilot Test RDD response rate was 18 percent, which was 50 percentage points below the NCVS in the same CBSAs. While the RSA Pilot Test did not find non-response bias to be an issue with the CATI survey, this conclusion is based on assumptions that could not be fully tested. In addition, the response rates for RDD continue to decline over time, which may introduce non-response bias and, at the very least, make the survey more expensive to implement. Finally, an RDD frame does not allow an easy way to ask respondents to move to the web.

An independent ABS sample, which contacts respondents in-person, should be able to achieve comparable response rates to the NCVS. However, based on results from this study, the cost of conducting the survey would be approximately 4 times higher than an independent RDD survey (Chapter 13).

When designing the separate survey, an initial cost assessment should be completed to make a more concrete comparison of the two methods of sampling. Several different designs should be specified (e.g., oversampling parameters; mode of interview; reference period; sample size) and costs generated for doing the survey as part of the core NCVS and for different types of independent samples.

Recommendation 6: Define “rape” as penetration involving physical force or inability to consent. Define “sexual assault” as other types of kissing or other sexual contact involving physical force or inability to consent. Include attempts and threats for both rape and sexual assault.

Central elements of all U.S. state definitions of rape are the behaviors that are included (penetration of some type) and the use of force or an inability to give consent. The latter includes being unconscious or asleep. Many states also include provisions when the respondent is unable to give consent due to impairment because of the use of alcohol or drugs. We do not recommend including in this definition exposure or exploitative videos/photographs. These are not consistently

included under rape and sexual assault laws. They also made up a very small proportion of the incidents on the RSA Pilot Test.

The NCVS has traditionally included both threats and attempts for all types of predatory crimes, including rape and sexual assault. This is consistent with state laws. For this reason, we also recommend including threats and attempts in the definition as well. The RSA Pilot Test measures of attempts and threats were among the least reliable collected on the study. Respondents' views of an attempt or a threat tended to be broader than intended by the NCVS. Recommendations below make suggestions on how these measures could be improved.

Sexual assault should be based on a similar definition, which includes sexual contact involving force or inability to consent.

A separate set of indicators should be collected on unwanted sexual contact that involve non-physical threats, such threatening social or financial consequences or promising rewards. This form of coercion is an important form of unwanted sexual contact, which is included as sexual assault in some states.

Recommendation 7: The mode for the separate survey should be adapted to other features of the design, such as the sample design and methods to contact the respondent.

The results of the RSA Pilot Test did not find significant differences in the incidence and prevalence rates between the ACASI and CATI interviews for the general population sample. This suggests that the selection of mode of interview for the separate survey should be based on factors other than its effect on reporting rape and sexual assault. The one significant advantage of the ACASI design is that it is more conducive to collecting details about what happened during the incident. A number of the questions on the DIF relate to the circumstances that involve multiple response categories (e.g., victim-offender relationship; type of force or coercion used; reasons for not reporting to the police; type of injury suffered). For example, collecting information on the specific behaviors that occurred during the incident involves asking about different types of penetration and sexual touching, as well as distinguishing between completed and non-completed acts. The telephone survey was designed to use "yes/no" questions for each behavior to preserve confidentiality, while a self-administered survey has the advantage of presenting all of the possible options to the respondent, who can then pick the appropriate responses. As found on the RSA Pilot Test, this aspect of the design not only gives an ACASI design more flexibility with respect to

questionnaire design, but it also results in a shorter interview and lower rates of missing detailed incident forms.

A separate RSA survey that drew sample from the ongoing NCVS should maximize use of a self-administered survey, such as ACASI. This promotes the privacy of the interview in a household setting, as well as being less expensive to administer. To boost the response rate, non-respondents can be followed up with a telephone interview which takes the appropriate precautions to preserve privacy and confidentiality of the responses. If the separate RSA survey draws a sample independent of the ongoing NCVS, then the mode should be tailored to the specific design. For example, the telephone would naturally be used as part of an RDD sample, whereas ACASI would be implemented for an in-person contact or one that asks respondents to go to the web.

15.3.3 Length and Structure of a Separate Survey

In this section, recommendations are made with respect to the basic structure of a separate survey.

Recommendation 8: The separate survey should use behaviorally specific questions, a detailed incident form and a narrative.

This study found all three elements of the survey (BSQs, DIF and narrative) to be important contributors to the measurement of RSA. The RSA Pilot Test did not determine the optimum number of BSQs that are needed to fully measure rape and sexual assault. On the RSA Pilot Test, twelve BSQs were used to target the specific behaviors and tactics that made up the definition of rape and sexual assault. Three of these questions asked about exposure and forcing someone to take photos/videos. As noted above, these items had very low prevalence and did not contribute substantially to RSA estimates and could be dropped. When considering the BSQs for the survey, the remaining nine questions account for more than 90 percent of the incidents classified as rape and sexual assault. The six items targeting rape accounted for about 83 percent of these incidents reported on the survey. The remaining three items targeted sexual assault. These accounted for 75 percent of the sexual assaults. Some items targeting rapes led to reports of sexual assaults (21% of sexual assault) and those targeting sexual assaults led to reports of rapes (10% of rapes). These nine questions account for 93 percent of the rapes and 97 percent of the sexual assaults.

It may be possible to combine some of these questions to cover more than one type of behavior. For example, the recent Campus Climate Survey Validation Study (Krebs et al., 2016) had two multi-part questions, which resulted in prevalence rates equivalent to several other campus climate surveys (e.g., Cantor et al., 2016). On the other hand, the results of the reliability analysis for the RSA Pilot Test suggest that combining many items in a single question may lead to additional sources of variability in the estimates by making it more difficult for respondents to understand the questions. Combining questions may also reduce the number of individuals that report multiple incidents. Determining the number of questions needed could be addressed in the final development of the separate survey.

The RSA Pilot Test found the DIF added significantly to the information collected on the survey. For this reason, it is recommended to administer a DIF for each incident reported from the victimization screener. The DIF made it possible to describe victimization experiences at an incident level. The DIF was used to classify the event into a particular category of crime. It also provided details on the circumstances of the incident (e.g., location, victim–offender relationship, and consequences). For example, collecting data on each incident made it possible to describe events involving both force and alcohol/drug facilitation. Similarly, the DIF provided the information to describe tactics that involved by coercive or other tactics that do not necessarily involve force or alcohol/drug use (e.g., threats to work or financial resources).

The RSA Pilot Test found problems with some of the items used on the DIF. For example, analysis of consistencies with the screener items found the questions on unwanted behaviors were problematic for some respondents. Similarly, the questions on tactics related to sexual assault were missing several key actions (e.g., unexpected grabbing). Below, recommendations are made on how the items on the DIF can be improved.

It is also recommended that respondents be asked to provide a narrative at the end of each incident form. Respondents were generally willing to provide a narrative. This will make it possible to check the internal consistency of the information collected on both the DIF and the victimization screener by using the narrative as the most accurate description of the incident.

Recommendation 9: Limit the number of detailed incident forms to three. For respondents reporting more than three incidents, assign the highest priority to incidents involving penetration, followed by those involving sexual touching. Subsample incidents if a respondent has multiple incidents involving unwanted sexual contact.

The RSA Pilot Test limited the number of incidents forms to three. A significant number of RSA Pilot Test respondents reported multiple incidents occurring within the reference period. For example, about 25 percent of those reporting completed rape had two or more incidents in the past 12 months. Filling out multiple detailed incident forms adds to the length of the interview (e.g., see timings in Chapter 6). Some respondents became frustrated, especially on the telephone, when asked about the details for incidents that were very similar, especially incidents involving unwanted sexual kissing or groping, which in some cases were not particularly salient.

The analysis did not implement any procedures to estimate the number of incidents that were excluded because of capping the number of DIFs to three. For example, if a respondent reported three instances of forced penetration and a sexual assault, no DIF was requested for the assault. When capping incidents at three, it will be necessary to develop an estimation procedure, using the responses to the screener, to impute for the incidents that were not collected because they were too low priority for individuals reporting multiple incidents.

The survey did not institute any special procedures for serial experiences in which the victim has problems defining the boundaries of an incident. On the NCVS, these are called series incidents. If respondents cannot remember the details of a particular incident, they are asked to provide information about the most recent one. A similar procedure should be considered for the proposed supplement survey. Respondents reporting multiple incidents of the same type should be asked if they can remember the details for each incident. If they can, respondents should fill out a DIF for each one up to the limit of three. If the details cannot be remembered, the respondent can be asked to report for the most recent time it happened and then asked about any other incidents that had been reported

15.3.4 Measurement of Key Concepts

In this section, recommendations are made with respect to the measurement of key concepts on the victimization screener, the detailed incident form, and the classification scheme.

Recommendation 10: If the separate survey cannot use the same bounding procedure as the NCVS, use a sequence of questions that first ask about lifetime and then ask for anything occurring in the last 12 months.

The NCVS currently bounds most interviews by using the previous interview as a cognitive marker for the beginning of the reference period. Bounding does not occur for the very first visit to

the household. The data from this visit are adjusted to account for external telescoping. The RSA Pilot Test was a one-time survey that could not use the same procedure. To reduce the effects of external telescoping, two procedures were used. As discussed in section 15.1.4, the internal bounding procedure used on the CATI survey seemed to be effective at reducing external telescoping.

If the separate survey is conducted as part of a supplement to the NCVS, then the previous NCVS interview can be used as the boundary for the beginning of a 6-month reference period. However, if it is desirable to use a one-year reference period, or if an independent sample is used for the rape and sexual assault survey, then the two-step process asking for lifetime and then the target reference period (e.g., 12 or 6 month) is advisable based on the findings from this research.

Recommendation 11: Revise key RSA Pilot Test victimization screening items such that (1) when describing penetration the wording should communicate that very slight penetration should be counted and (2) non-consent is described as “against your will” for items involving physical force.

The first part of this recommendation stems from reviewing narratives where respondents did not report penetration when the perpetrator stopped shortly after penetration occurred. Examples of suggested wording to several of the RSA Pilot Test screening questions are presented below (note that underlined words reflect recommended wording changes from that used in RSA Pilot Test).

“Has a male ever used force or threats of force to make you have vaginal sex against your will? By vaginal sex, it means putting his penis, even if only slightly, in your vagina.”

“.....has anyone, male or female, used force or threats of force to make you have oral sex against your will? By oral sex, it means that someone penetrated your vagina or anus, even if only slightly, with their mouth or tongue.....”

“.....has a male used force or threats of force to make you have anal sex against your will? By anal sex, it means that a man or boy put his penis, even if only slightly, in your anus.”

“.....has a male or female used force or threats of force to put fingers or a foreign object in your vagina or anus, even if only slightly, against your will?”

The second recommendation is to use the phrase “against your will” as the primary way to communicate non-consent in the questions that are intended to ask about physical force. Other terms that could be used are “without consent,” “when you did not want to do it,” or “unwanted.” Respondents to the cognitive interviews found these other terms to be too broad and subject to

different interpretations. The phrase “against your will” is consistent with statutory language that was intended to convey incidents involving force, as well as BJS’s current definition (Truman & Morgan, 2016). One possible drawback to using the phrase “against your will” is that it is too narrow. It may not capture unwanted events where the respondent is either unable to consent or is coerced into the behavior. The RSA Pilot Test included questions on being unable to consent, as well as other types of non-physical tactics.

Recommendation 12: Simplify the questions on unwanted behavior on the detailed incident form.

The questions on the DIF that asked about unwanted behavior (D1a – D1d; D2a – D2e) were found to be problematic for some respondents. These items had the highest levels of missing data on the ACASI interview. Behavior coding of a sample of CATI interviews found that interviewers had a hard time administering these questions. There were several instances in which the responses to these items were not consistent with the narratives and these items had relatively low reliabilities. Some of these problems were due to the structure of the items. An example of the first question on unwanted penetration is shown below:

Did the person threaten to, try to, or actually put his penis in your vagina when you didn't want it to happen? (Mark one response)

Yes, the person verbally threatened to do this but did not physically try to do it

Yes, the person physically tried to do this but did not actually do it

Yes, the person actually did do this

No, this did not happen

The question asked about four different conditions (threatened, attempted, completed, did not happen). This format was used to ask about all of the different unwanted behaviors covered by the survey, including oral, digital and anal penetration, unwanted kissing, groping, attempted kissing and groping, exposure, and making unwanted videos.

Some of the problems with the question are related to asking respondents to make multiple distinctions between somewhat subtle differences (e.g., attempted vs threatened). The question is long and respondents may not remember which option is applicable once the question is read. Redesign of this question should consider first asking if any of the unwanted behaviors were completed at all. A follow-up question would then be presented asking if there were any attempts or

threats to do any of the behaviors. An example of this sequence is provided below, after other aspects of asking about unwanted behaviors are discussed. (See exhibit 15-2.)

Recommendation 13: Collect a combined measure of threats and attempts. Include follow-up questions that operationalize how the threat or attempt was carried out for unwanted penetration.

Much of the inconsistency associated with measuring unwanted behaviors was related to attempted or threatened behavior. The reliability of the responses for threats and attempts from the re-interview analysis was particularly low (see Chapter 12). As noted in Chapter 7, respondents varied how they interpreted what constituted an attempt or threat of unwanted penetration. The analysis of the narratives, in particular, found a number of inconsistencies. In some instances, attempted or completed unwanted sexual touching or kissing was seen as an attempt at penetration, even though the narratives showed no indications of physical or verbal actions indicating that penetrative behavior was imminent. From a legal perspective, an attempt or threat needs to have both the intent and capacity for the act to be carried out. This makes it difficult to measure on a survey because it is necessary to measure the context of the event. For example, verbal harassment in a public space (e.g., walking down the street) would not be considered a credible threat of rape from a legal perspective, but could reasonably be considered a threat by the victim.

Some consideration should be given to whether the NCVS wants to measure threats and attempts, given the difficulties in the measurement. Below we propose a possible sequence of questions that might be asked to further operationalize the measurement. But the burden placed on the respondent should be weighed against the proposed approach. Further development of these items, should be done to assess whether these are options that should be considered.

The NCVS operationalizes threats and attempts by asking follow-up questions that ask what constituted the threat or attempt. A similar approach could be taken for measuring rape and sexual assault. We recommend using the sequence below for those who report attempted or threatened unwanted penetration. We also recommend not asking this level of detail for attempted or threatened sexual contact. This minimizes the burden on the respondent and reduces the categorization of attempted/threatened rape for incidents that are more appropriately classified as sexual assault or attempted sexual assault.

The approach would first ask if there were any attempts or threats to carry out the unwanted behavior. This would be asked for all unwanted behaviors.

....did the person verbally THREATEN or physically TRY to (BEHAVIOR), but did not actually do it?

Yes

No

A response of “yes” to this item, for non-penetrative sexual contact, would be classified as a sexual assault. If the respondent reports a threat or attempt at some type of penetration, the respondent would be asked what constituted the threat or attempt:

- a. *by trying to physically force you to do it*
- b. *by threatening to physically force you to do it*
- c. *by threatening to hurt someone else unless you did it*
- d. *by trying to do it while you were asleep, unconscious or unable to consent because you were high on alcohol or drugs.*
- e. *(ask if a-d = no) by verbally telling you they wanted to do it*
- f. *(ask if a-d = no) by doing something else to you, like kissing, touching, or grabbing you, that made it seem like they wanted more.*

If any of these options are selected, the incident would be considered an attempted or threatened rape.

Some of these options do not fully measure the credibility of the threat. For example, verbal threats may be made in situations where it is not credible (e.g., see example of sexual harassment in a public place). An additional measure could be to ask the respondents who only select options b, e, or f:

At the time, were you afraid that the person was actually going to (BEHAVIOR)?

This is an approach that is currently being used to measure stalking, which relies on self-reports from victims about fear of being unsafe.

Recommendation 14: Link the force questions on the detailed incident form to the type of unwanted behavior that is reported. Separate the tactic questions for penetration from other types of sexual contact.

The RSA Pilot Test had separate questions on behaviors and tactics. A respondent could report several unwanted behaviors, as well as several tactics, but there was no linkage between the two. It is recommended that the behavior and tactic questions be interwoven so there is a direct linkage between the two. The initial sequence would ask about each type of unwanted behavior:

For these questions please only focus on the parts of the incident that you did not want to happen. In this particular incident... (Mark yes or no for each)

- a. *Was there unwanted vaginal sex?*
- b. *Was there unwanted oral sex?*
- c. *Was there unwanted anal sex?*
- d. *Was there unwanted sexual penetration with a finger or object in your [FEMALE: vagina or] anus or in the person's vagina or anus?*
- e. *Was there any unwanted sexual touching, such as kissing, touching your sexual body parts, or grabbing, fondling, or rubbing up against you in a sexual way, even if it was over your clothes?*

These items do not repeat the definitions of the behaviors, as they were defined in the screening instrument. If there is a yes to any type of penetration (options a – d above) then ask:

You said that there was (READ ALL YES RESPONSES FROM #Q1: oral sex, anal sex, sexual intercourse, sexual penetration with a finger or object) when you did not want it to happen. In this particular incident, ... (Mark yes or no for each)

- a. *Did the person use physical force, such as holding or pinning you, hitting or kicking you, or using a weapon?*
- b. *Did the person threaten to hurt you or someone close to you?*
- c. *Did this happen while you were blacked out, unconscious, or asleep?*
- d. *Did this happen while you were unable to consent because you were too drunk or high?*
- e. *Did the person use some other type of pressure, such as threatening to cut off financial support, threatening to cause problems for you, or promising rewards?*
- f. *[IF NO TO ALL IN 2A-E] Did this happen for some other reason (specify)?*

If the incident involved unwanted sexual touching, the respondent would be asked,

You said that some kind of sexual touching occurred when you did not want it to happen. In this particular incident, ... (Mark yes or no for each)

- a. *Did the person forcibly kiss you, grab, or touch your sexual body parts?*
- b. *Did the person use physical force, such as holding or pinning you, hitting or kicking you, or using a weapon?*
- c. *Did the person threaten to hurt you or someone close to you?*
- d. *Did this happen while you were blacked out, unconscious, or asleep?*
- e. *Did this happen while you were unable to consent because you were too drunk or high?*
- f. Did the person use some other type of force or pressure, such as threatening to cut off financial support, threatening to cause problems for you, or promising rewards?
- g. *[IF NO TO ALL IN 3A-F] Did this happen for some other reason (specify)?*

Option 'a' above was added as a type of force unique to sexual contact. This was commonly written in on the RSA Pilot Test, which did not include this option.

Exhibit 15-2 provides the full sequence of questions for the unwanted behavior questions.

Recommendation 15: Consider using multiple criteria when classifying an incident as being related to alcohol and drugs.

Classification of the incident as due to inability to consent because of alcohol or drugs was based on the response to one of three questions—(1) she was unable to consent (question G12a), (2) she was passed out for part of the incident (question G10) or (3) she was passed out for all of the incident (G10). As noted above, slightly more than half of the incidents that were classified as being unable to consent were from responses to the direct question (G12a). The remaining were classified in this category because they were unconscious for at least part of the incident.

Overall, the above discussion on false positives provides evidence that the measures used to define when the victim is unable to consent have face validity. Most who said they were unable to consent met some criteria of being intoxicated to the point that they had trouble making decisions and provided some indication that another individual would recognize the victim was not in condition to consent. Of course, the indicator is subject to some error with both false positives and

false negatives. The phrase “unable to consent” is open to some interpretation and likely contributes to this error. We recommend including the same, or related, indicators of drug and alcohol use that were measured on the RSA Pilot Test on the separate survey to give analysts the ability to see if results differ by how inability to consent is defined.

15.4 Summary and Limitations

This pilot study provides a blueprint for collecting rape and sexual assault data within the ongoing NCVS program. The proposed revisions should significantly improve the measure of rape and sexual assault over the current NCVS. This report provides evidence that the improved measures meet most legal definitions of these crimes and are correlated with a number of socio-demographic characteristics that are expected to be related to risk of these types of crimes. As part of the NCVS redesign, it is recommended that the behavior-specific screening questions and revisions to the current detailed incident form expand the types of incidents that are measured on the survey. Comparison of the RSA Pilot Test and NCVS rates of rape and sexual assault suggest that much of the difference between the two is due to differences in the scope, content, and context of the two surveys. Revising the questions targeting rape and sexual assault on the NCVS should improve the measures that are collected.

If the redesign of the NCVS does not fully address the measurement issues, it is recommended that a separate survey on rape and sexual assault be conducted. This separate survey could sample from respondents to the ongoing NCVS or draw an independent sample. Cost data should be used to map out these differences in more detail. The separate survey should adopt many of the features of the RSA Pilot Test, with modifications to improve the measurement of attempted and completed crimes.

This study has several limitations. One is that the sample was restricted in several ways. First, it did not include males in the sample. As noted in Chapter 2, this exclusion was done for purely cost reasons. It was not possible to collect enough data for both sexes, given the very low rates of rape and sexual assault for males. We do not believe this is a major restriction. The survey language needs to be adapted and tested with men. Second, the analysis in this report did not concentrate on anyone over the age of 49. This was done to ensure enough reports of rape and sexual assault to test the full design. Those age 50 and over have very low annual rates of rape and sexual assault, as confirmed by the CATT survey. Similarly, the study was conducted in five large CBSAs and did not

include less densely populated areas. While there is no reason to believe there are unique effects of the methodologies implemented on RSA Pilot Test for older adults or residents of these particular cities, future development work should expand to an older age range and a larger geographic universe.

Perhaps a more constraining limitation of the sample is that it did not include any children under that age of 18. The NCVS includes children age 12 to 17 and provides a wide range of estimates for this age group. According to the NCVS, this age group has the highest rates of rape and sexual assault (Planty et al., 2013). Considerable development work is needed to adapt the RSA Pilot Test methods to this age group. It is unclear how difficult it might be to get parental consent for an interview that included the types of questions on the RSA Pilot Test. Elevating the profile of RSA incidents on the ongoing NCVS or conducting a more specialized survey on the topic also raises questions about how to handle reports of RSA that would normally be required to be reported to state authorities (i.e., mandatory reports). Finally, the language used on the RSA Pilot Test was not tested with children. While BJS has some experience with this type of survey with incarcerated youth (e.g., Beck et al., 2013), this is a very specialized group and cannot be easily extrapolated to the general population.

A second limitation is that the overall response rate was below what could be achieved on the NCVS. The analysis of nonresponse did not find significant nonresponse bias when comparing the 18 percent response rate for the CATI and the 40 percent response rate for the ACASI. Nonetheless, the NCVS response rate was 68 percent in these particular CBSAs. The effects of nonresponse on the estimates cannot be ruled out when comparing to the NCVS.

Finally, a third limitation is that this study was conducted as a one-time survey, conducted by a private contractor. The NCVS is conducted by the Census Bureau, an agency of the Federal government, as part of a rotating panel design. It will be important to examine the recommended procedures within the context of the NCVS framework and Census Bureau operations when finalizing the methodology for measuring sexual victimization. As noted above, there are a number of differences between the rotating panel design of the core NCVS and those used on the RSA Pilot Test. If prior methodological studies are any indication (e.g., Biderman et al., 1986), these differences can have a significant effect on the incidence and prevalence rates produced by the ongoing NCVS.

Exhibit 15-2. Example of revised sequence to measure behaviors and tactics

- Q1.** **For these questions, please only focus on the parts of the incident that you did not want to happen. In this particular incident, ... (Mark yes or no for each)**
- a. Was there unwanted vaginal sex?
 - b. Was there unwanted oral sex?
 - c. Was there unwanted anal sex?
 - d. Was there unwanted sexual penetration with a finger or object in your [FEMALE: vagina or] anus or in the person's vagina or anus?
 - e. Was there any unwanted sexual touching, such as kissing, touching your sexual body parts, or grabbing, fondling, or rubbing up against you in a sexual way, even if it was over your clothes?
- Q2.** **(If any in Q1a-Q1d=yes) You said that there was (READ ALL YES RESPONSES FROM #Q1: oral sex, anal sex, sexual intercourse, sexual penetration with a finger or object) when you did not want it to happen. In this particular incident, ... (Mark yes or no for each)**
- a. Did the person use physical force, such as holding or pinning you, hitting or kicking you, or using a weapon?
 - b. Did the person threaten to hurt you or someone close to you?
 - c. Did this happen while you were blacked out, unconscious, or asleep?
 - d. Did this happen while you were unable to consent because you were too drunk or high?
 - e. Did the person use some other type of pressure, such as threatening to cut off financial support, threatening to cause problems for you, or promising rewards?
 - f. [IF NO TO ALL IN 2A-E] Did this happen for some other reason (specify)?
- Q3.** **(If Q1e=yes) You said that some kind of sexual touching occurred when you did not want it to happen. In this particular incident, ... (Mark yes or no for each)**
- a. Did the person forcibly kiss you, grab or touch your sexual body parts?
 - b. Did the person use physical force, such as holding or pinning you, hitting or kicking you, or using a weapon?
 - c. Did the person threaten to hurt you or someone close to you?
 - d. Did this happen while you were blacked out, unconscious or asleep?
 - e. Did this happen while you were unable to consent because you were too drunk or high?
 - f. Did the person use some other type of force or pressure, such as threatening to cut off financial support, threatening to cause problems for you, or promising rewards?
 - g. [IF NOT TO ALL IN 3A-F] Did this happen for some other reason (specify)?

Exhibit 15-2. Example of revised sequence to measure behaviors and tactics (continued)

- Q4.** (If all Q1a – Q1d = no then ask items 4a-d; If Q1e = no, ask item 4e). Did the person verbally THREATEN or physically TRY to (BEHAVIOR), BUT DID NOT ACTUALLY DO IT?
- a. Have vaginal sex with you
 - b. Have oral sex with you
 - c. Have anal sex with you
 - d. Sexually penetrate with a finger or object in your [FEMALE: vagina or] anus or make you put your finger or an object in the person's vagina or anus?
 - e. Kiss, touch your private parts, or grab, fondle, or rub up against you in a sexual way, even if it was over your clothes?
- Q5.** (If any of Q4a – Q4d=yes) You said that the person threatened to or tried to (READ ALL YES RESPONSES FROM #Q4: have oral sex, have anal sex, have sexual intercourse, sexually penetrate you with a finger or object) when you did not want it to happen. How did the person threaten or attempt this? Did they...
- a. Try to physically force you to do it
 - b. Threaten to physically force you to do it
 - c. Threaten to hurt someone else unless you did it
 - d. Try to do it while you were asleep, unconscious, or unable to consent because you were high on alcohol or drugs.
 - e. (ask if all in a – d = no) Verbally tell you they wanted to do it
 - f. (ask if all in a-d = no) Do something else to you, like kissing, touching, or grabbing you, that made it seem like they wanted more
- Q6.** (If Q5b or Q5e = yes or Q5f = yes) At the time, were you afraid that the person was actually going to (BEHAVIOR)?

Source: Bureau of Justice Statistics, Rape and Sexual Assault (RSA) Pilot Test, 2014-2015.

References

- American Association for Public Opinion Research. (2016). *Standards definitions. Final dispositions of case codes and outcome rates for surveys.* Accessed on November 16, 2016, at http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.
- Anderson, J., & Powers, P. (2016). Alcohol-facilitated sexual assault: Who needs force when you have alcohol? *AEquitas: The Prosecutors' Resource for Violence Against Women*. Last accessed November 17, 2016 at <http://www.aequitasresource.org/trainingDetail.cfm?id=132>
- Beck, A.J., Cantor, D., Hartge, J., & Smith, T. (2013). Sexual victimization in juvenile facilities reported by youth, 2012 (NCJ 241708). Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.
- Berzofsky, M., & I. Carrillo-Garcia (2017). Determining the Optimal Number of Interview Waves in a Panel Survey with Application to the NCVS: Evaluation and Recommendations. Research Report, Bureau of Justice Statistics, NCJ 249878, <http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5825>
- Bialik, C. (2013). Statistics shine little light on rape rates. *The Wall Street Journal*.
- Black, M.C., Basile, K.C., Breiding, M.J., Smith, S.G., Walters, M.L., Merrick, M.T., Chen, J., & Stevens, M.R. (2011). *The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 Summary Report*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.
- Black, M.C., Kresnow, M., Simon, T.R., Arias, I., & Shelley, G. (2006). Telephone survey respondents' reactions to questions regarding interpersonal violence. *Violence and Victims*, 21(4), 445-459.
- Blumberg, S.J., & Luke, J.V. (2014). *Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2013*. Centers for Disease Control and Prevention.
- Blumberg, S.J., & Luke, J.V. (2016). *Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2015*. Centers for Disease Control and Prevention.
- Boyle, J., Bucuvalas, M., Piekarski, L., & Weiss, A. (2009). Zero banks: Coverage error and bias in RDD samples, based on hundred banks with listed number. *Public Opinion Quarterly*, 73(4), 729–750.
- Brick, J. M., & Williams, D. (2013). Explaining rising nonresponse rates in cross-sectional surveys. *Annals of the American Academy of Political and Social Science*, 645(1), 36–59.
- Brick, J.M., Cervantes, I.F., Lee, S., & Norman, G. (2011). Nonsampling errors in dual frame telephone surveys. *Survey Methodology*, 37(1), 1-12.

- Brick, J.M., Rizzo, L., & Wernimont, J. (1997, June). *Reinterview results for the school safety and discipline and school readiness components* (NCES 97339). Washington, DC: U.S. Department of Education, National Center for Education Statistics. [Available as a PDF at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=97339>]
- Cantor, D. (1989). Substantive implications of selected operational longitudinal design features: The National Crime Survey as a case study. In D. Kasprzyk & G. Duncan (Eds.), *Panel surveys* (pp. 1-30). New York, NY: John Wiley.
- Cantor, D. (2009). Two approaches to address non-response: A case study with the Health Information National Trends Survey (HINTS). In L. Rutten, B. Hesse, R. Moser, and G. Kreps (Eds.), *Health communication*. Cresskill, NJ: Hampton Press.
- Cantor, D., & Williams, D. (2013). *Assessing interactive voice response for the National Crime Victimization Survey*. Final report for the Bureau of Justice Statistics. Prepared under Contract 2008-BJ-CX-K066. Accessed November 11, 2016 at <http://www.bjs.gov/content/pub/pdf/Assessing%20IVR%20for%20the%20NCVS.pdf>
- Cantor, D., Steiger, D., Townsend, R., Catalano, S., & Ricci, K. (2014). *Using cognitive interviews to test behaviorally specific questions on rape*. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Anaheim, CA.
- Cantor, D., Townsend, R., & Sun, H. (2016). *Methodology report for the AAU Campus Climate Survey on Sexual Assault and Sexual Misconduct* (for Association of American Universities). Rockville, MD: Westat.
- Catalano, S. (2016). *Interviewing conditions in the National Crime Victimization Survey, 1993–2013*. Bureau of Justice Statistics, U.S. Department of Justice, NCJ 249682.
- Coker, A.L., & Stasny, E.A. (1994). *Adjusting the National Crime Victimization Survey's estimates of rape and domestic violence for "gag" factors, final report*. NCJ 173061, Washington, DC: United States Department of Justice, National Institute of Justice.
- Cook, S.L., Gidycz, C.A., Koss, M.P., & Murphy, M. (2011). Emerging issues in the measurement of rape victimization. *Violence Against Women*, 17, 201-218.
- Couzens, L.G., Berzofsky, M. and C. Krebs. (2014). Analyzing potential mode effects in the National Crime Victimization Survey. *Proceedings of the Survey Research Methods Section, American Statistical Association*. <http://ww2.amstat.org/sections/srms/Proceedings/>
- Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse error over the past quarter century. *Public Opinion Quarterly*, 69, 87–98.
- Decker, S.E., Naugle, A.E., Carter-Visscher, R., Bell, K., & Seifert, A. (2011). Ethical issues in research on sensitive topics: Participants' experiences of distress and benefit. *Journal of Empirical Research on Human Research Ethics*, 6(3), 55-64. doi: 10.1525/jer.2011.6.3.55

- DePrince, A.P., & Chu, A. (2008). Perceived benefits in trauma research: Examining methodological and individual difference factors in responses to research participation. *Journal of Empirical Research on Human Research Ethics*, 3(1), 35-47. doi: 10.1525/jer.2008.3.1.35
- DiLillo, D., DeGue, S., Kras, A., Di Loreto-Colgan, A. R., & Nash, C. (2006). Participant responses to retrospective surveys of child maltreatment: Does mode of assessment matter? *Violence and Victims*, 21(4), 410-424.
- Dutwin, D., & Malarek, D. (2014). The use of recent activity flags to improve cellular telephone efficiency. *Survey Practice*, 7. Retrieved Aug. 10, 2016, from <http://www.surveypartice.org/index.php/SurveyPractice/rt/printefriendly/57/html>.
- Fahimi, M., Kulp, D., & Brick, J.M. (2009). A reassessment of list-assisted RDD methodology. *Public Opinion Quarterly* 73(4), 751-760.
- Finch, J. (1987). The vignette technique in survey research. *Sociology*, 21, 105-114.
- Fisher, B.S. (2009). The effects of survey question wording on rape estimates: Evidence from a quasi-experimental design. *Violence Against Women*, 15, 133–147.
- Fisher, B., Cullen, F.T., & Turner, M.G. (2000). *The sexual victimization of college women*. National Institute of Justice, U.S. Department of Justice, NCJ 182369.
- Fisher, B., & Cullen, F. (2000). Measuring the sexual victimization of women: Evolution, current controversies and future research. In National Institute of Justice (Ed.), *Measurement and Analysis of Crime and Justice*, Vol. 4, National Institute of Justice, Washington DC.
- Fowler, Jr, F.J., & Mangione, T.W. (1990). *Standardized survey interviewing: Minimizing interviewer-related error* (Vol. 18). Newbury Park, CA: Sage.
- GAO. (2016). *Sexual violence data: Actions needed to improve clarity and address differences across federal data collection efforts*. GAO-16-546. Government Accountability Office, US Government Printing Office, Washington DC.
- Gilbert, N. (1997). Advocacy research and social policy. In *Crime and justice: A review of research*, Michael Tonry (ED). Vol. 22. Chicago: University of Chicago Press.
- Graham, D. (1977). Response errors in the National Crime Survey: July 1974-June 1976. U.S. Bureau of the Census memorandum, August 10, 1977. In R.G. Lehnert & W. G. Skogan (Eds.), *The National Crime Survey: Working Papers Volume 11: Methodological studies* (pp. 58-64). Washington, DC: Bureau of Justice Statistics.
- Groves, R., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Siknger, E., & Tourangeau, R. (2004). *Survey methodology*. New York: Wiley.
- Groves, R.M., & Magilavy, L.J. (1986). Measuring and explaining interviewer effects in centralized telephone facilities. *Public Opinion Quarterly*, 50(2), 251-266.

- Groves, R.M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70, 46–75.
- Humphreys, T. (2007). Perceptions of sexual consent: The impact of relationship history and gender. *Journal of Sex Research*, 44(4), 307-315.
- Jaffe, A.E., DiLillo, D., Hoffman, L., Haikalis, M., & Dykstra, R.E. (2015). Does it hurt to ask? A meta-analysis of participant reactions to trauma research. *Clinical Psychology Review*, 40, 40-56.
- Keeter, S., Kennedy, C., Dimock, M., Best, J. & Craighill, P. (2006). Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. *Public Opinion Quarterly*, 70,759–79.
- Keeter, S., Miller, C., Kohut, A., Groves, R.M. & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly*, 64,125–48.
- Kilpatrick, D. (2004). What is violence against women? Defining and measuring the problem. *Journal of Interpersonal Violence*, 19, 1209-1234.
- Kilpatrick, D., Edmunds, C., & Seymour, A. (1992). *Rape in America: A report to the nation*. Arlington, VA: National Victim Center and Medical University of South Carolina.
- Kilpatrick, D., Resnick, H., Rugiero, K., Conoscenti, L., & McCauley, J. (2007). *Drug-facilitated, incapacitated, and forcible rape: A national study*. Charleston, SC: Medical University of South Carolina and National Crime Victims Research and Treatment Center.
- Koss, M.P. (1988). Hidden rape: sexual aggression and victimization in a national sample of students in higher education. In A. W. Burgess (Ed.), *Rape and sexual assault. Volume 2*, pp. 3-25. New York: Garland.
- Koss, M.P. (1993). Detecting the scope of rape: A review of prevalence research methods. *Journal of Interpersonal Violence*, 8(2), 198-222.
- Koss, M.P., & Gidycz, C.A. (1985). Sexual experiences survey: Reliability and validity. *Journal of Consulting and Clinical Psychology*, 53(3), 422.
- Koss, M.P., Gidycz, C.A. & Wisneiewski, N. (1987). The scope of rape: incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. *Journal of Consulting and Clinical Psychology*, 55, 162-170.
- Krebs, C., Lindquist, C., Berzofsky, M., Shook-Sa, B., Peterson, K., Planty, M., Langton, L., & Stroop, J. (2016). *Campus climate survey validation study final technical report*. Bureau of Justice Statistics, U.S. Department of Justice, R&DP-2015:04, NCJ 249545.
- Krebs, Christopher P., Christine H. Lindquist, Tara D. Warner, Bonnie S. Fisher, Sandra L. Martin, & Childers, J. (2011. Comparing sexual assault prevalence estimates obtained with direct and indirect questioning techniques. *Violence Against Women*, 17(2), 219-235.

- Kroh, M., Winter, F., & J. Schupp. (2016). Using person-fit measures to assess the impact of panel conditioning on reliability. *Public Opinion Quarterly* 80: 914-942.
- Kruttschnitt, C., Kalsbeek, W.D., & House, C.C. (2014). *Estimating the incidence of rape and sexual assault*. Washington, DC: National Academies Press.
- Laaksonen, S., & Heiskanen, M. (2014). Comparison of three modes for a crime victimization survey. *Journal of Survey Statistics and Methodology*, 2, 459-483.
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Lauritsen, J. (2012). *Subpopulations at high risk for rape and sexual assault: What does the NCVS tell us?* Report prepared for the Panel on Measuring Rape and Sexual Assault in Bureau of Justice Statistics Surveys, Committee on National Statistics, Washington, DC.
- Lauritsen, J.L., & White, N. (2014). *Seasonal patterns in criminal victimization trends*. Bureau of Justice Statistics, June 2014 NCJ 245959.
- Lim, G.Y., & Roloff, M.E. (1999). Attributing sexual consent. *Journal of Applied Communication Research*, 27, 1-23.
- Lin, I.-F., & Schaeffer, N.C. (1995). Using survey participants to estimate the impact of nonparticipation. *Public Opinion Quarterly*, 59(2), 236–258.
- Livingston, J.A., Buddie, A.M., Testa, M., & VanZile-Tamsen, C. (2004). The role of sexual precedence in verbal sexual coercion. *Psychology of Women Quarterly*, 28, 287–297.
- Loftus, E.F., Klinger, M.R., Smith, K.D., & Fieldler, J. (1990). A tale of two questions: Benefits of asking more than one question. *Public Opinion Quarterly*, 54, 330-345.
- Lynch, J.P. (1996). Review: Clarifying divergent estimates of rape from two national surveys. *The Public Opinion Quarterly*, 60(3), 410-430.
- McClinton Appollis, T., Lund, C., de Vries, P. J., & Matthews, C. (2015). Adolescents' and adults' experiences of being surveyed about violence and abuse: A systematic review of harms, benefits, and regrets. *American Journal of Public Health*, 105(2), e31-45.
- Merkle, D., & Edelman, M. (2002). Nonresponse in exit polls: A comprehensive analysis, in Groves, R., Dillman, D., Eltinge, J. and R. Little (eds), *Survey nonresponse*, pp. 243 – 258, New York: Wiley.
- Montaquila, J.M., Brick, J.M., Williams, D. & Kim, K. (2013). A study of two-phase mail survey data collection methods. *Journal of Survey Statistics and Methodology* 1, 66–87.
- Newman, E., Risch, E., & Kassam-Adams, K. (2006). Ethical issues in trauma-related research: A review. *Journal of Empirical Research on Human Research Ethics*, 1(3), 29-46.

- Olson, K. (2006). Survey participation, nonresponse bias, measurement error bias and total bias. *Public Opinion Quarterly*, 70(5), 737-758.
- Peytchev, A., Baxter, R.K., & Carley-Baxter, L.R. (2009). Not all survey effort is equal. Reduction of nonresponse bias and nonresponse error. *Public Opinion Quarterly* 73,(4), 785-806.
- Planty, M., Langton, L., Krebs, C., Berzofsky, M., & Smiley-McDonald, H. (2013). *Female victims of sexual violence, 1994 – 2010*. Washington, DC: Bureau of Justice Statistics, NCJ 240655.
- Rand, M.R., & Rennison, C.M. (2005). Bigger is not necessarily better: An analysis of violence against women estimates from the National Crime Victimization Survey and the National Violence Against Women Survey. *Journal of Quantitative Criminology*, 21(3), 267-291.
- Rizzo, L., Brick, J.M., & Park, I. (2004). A minimally intrusive method for sampling persons in random digit dial surveys. *Public Opinion Quarterly*, 68(2), 267-274.
- Singer, P., & Ennis, S.R. (2003, September 24). *Census 2000 Content Reinterview Survey: Accuracy of data for selected population and housing characteristics as measured by reinterview* (Census 2000 Evaluation B.5). Washington, DC: U.S. Census Bureau. [Available at http://www.census.gov/pred/www/eval_top_rpts.htm]
- Sinozich, S., & Langton, L. (2014). *Rape and sexual assault victimization among college-age females, 1995 – 2013*. Washington, DC: Bureau of Justice Statistics, NCJ 248471
- Snijders, T.B., & Bosker, R.R.(1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, U.K.: Sage.
- Steiger, D., Cantor, D., Maitland, & Stapleton, M. (2014). *Reliability in qualitative research: Now what does that question mean to you?* Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Anaheim, CA.
- Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. (2014, September 4). *The NSDUH report: Substance use and mental health estimates from the 2013 National Survey on Drug Use and Health: Overview of findings*. Rockville, MD.
- Substance Abuse and Mental Health Services Administration. (2010). *Reliability of key measures in the National Survey on Drug Use and Health* (Office of Applied Studies, Methodology Series M-8, HHS Publication No. SMA 09-4425). Rockville, MD.
- Sudman, S., Finn, A., & Lannom, L. (1984). The use of bounded recall procedures in single interviews. *Public Opinion Quarterly*, 48, 520-524.
- Testa, M., VanZile-Tamsen, C., Livingston, J.A., & Koss, M.P. (2004). Assessing women's experiences of sexual aggression using the Sexual Experiences Survey: Evidence for validity and implications for research. *Psychology of Women Quarterly*, 29, 345-352.
- Tjaden, P., & Thoennes, N. (1998). *Stalking in America: Findings from the National Violence Against Women Survey*. Research in Brief, NCJ 169592. Washington, DC: U.S. Department of Justice.

- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133, 859-833.
- Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. New York, NY: Cambridge University Press.
- Tracy, C.E., Fromson, T.L., Long, J.G., & Whitman, C. (2013). *Statutory compilation rape and sexual assault laws*. Washington DC: AEQUITAS. Paper presented for the workshop convened by the Committee on National Statistics, June 5-7, Washington DC.
- Truman, J. (2011). *Criminal victimization, 2010*. Bureau of Justice Statistics, NCJ 234408.
- Truman, J.L., & Morgan, R.E. (2016). *Criminal victimization, 2015*. Bureau of Justice Statistics, U.S. Department of Justice, NCJ 250180.
- Tucker, C., Lepkowski, J.M., & Piekarski, L. (2002). The current efficiency of list-assisted telephone sampling designs. *Public Opinion Quarterly*, 66(3), 321-338.
- U.S. Department of Justice. (2011). *Methodological research to support the National Crime Victimization Survey: Self-report data on rape and sexual assault—pilot test solicitation*. Grants.gov number 2011-BJS-3011, p. 10, available at http://www.bjs.gov/content/pub/pdf/cvssdrsapt_sol.pdf
- Valpied, J., Cini, A., O'Doherty, L., Taket, A., & Hegarty, K. (2014). "Sometimes cathartic. Sometimes quite raw": Benefit and harm in an intimate partner violence trial. *Aggression and Violent Behavior*, 19, 673-685. doi: 10.1016/j.avb.2014.09.005
- Wager, N.M. (2012). Respondents' experiences of completing a retrospective web-based, sexual trauma survey: Does a history of sexual victimization equate with risk of harm? *Violence and Victims*, 27(6), 991-1004.
- Walker, E.A., Newman, E., Koss, M., & Bernstein, D. (1997). Does the study of victimization revictimize the victims? *General Hospital Psychiatry*, 19, 403-410.