Index

A

A/B studies, 35, 54, 216–217
Abbott prescription drug labels, 271–279
Accessibility checking tools, 229
Accessibility data, 227–231
Accessibility Valet Demonstrator, 229
AccMonitor (HiSoftware), 229
Accuracy, in reporting decimal places, 26
ACSI. *See* American Customer Satisfaction Index
Add Trend Line option (Excel), 32
Adjusted Wald Method, 69
Affective Computing Research Group, 186–188
After-Scenario Questionnaire (ASQ), 129–130
  in IVR case study, 245
  in OneStart case study, 282
Aggregating task ratings, 137–138
Albert, W., 129, 131, 179
Alternative design comparisons, 55
American Customer Satisfaction Index (ACSI), 151, 153–155
American Institutes for Research, 147
Analysis of variance (ANOVA)
  binary successes, 67
  pharmacist performance case study, 277
  types, 30–31
Analyzing and reporting metrics
  binary success, 67–68
  card-sorting data, 218–228
  efficiency, 88–90
  errors, 84–86
  frequency of unique issues, 109–111
  frequency of participants per issue, 111–112
  issues by category, 112–113
  issues by task, 113
  learnability data, 94–96
  levels of success, 72–73
  positive issues, 114
  precision in, 26
  self-reported data, 127–128
  time-on-task data, 77–79
Anchor terms, in Likert scales, 124
Anonymity, in self-reported data collection, 127
ANOVA. *See* Analysis of variance
ANOVA: Single-Factor option, 30–31
Apple IIe Design Guidelines, 102
Artifact bias, 116
ASQ. *See* After-Scenario Questionnaire
Assistance factors, in levels of success, 71
Attractiveness area, in WAMMI, 151, 153
Attribute assessment, in self-reported metrics, 158–161
Automated studies, 103
Automated tools
  accessibility-checking, 229
  measuring time-on-task, 75
Awareness
  and comprehension, 163–165
  increasing, 52
Awareness–usefulness gaps, 165
Axes
  bar graphs, 36–37
  labeling, 35
  line graphs, 39–40
  MDS plots, 224
  scatterplots, 32, 40–41
  stacked bar graphs, 43

B

Bailey, B. P., 181–182
Bailey, Robert W., 108, 252, 262
Baker, J., 88
Ball, Harvey, 205
Ballots
  comparisons, 148–150
  errors in, 6, 82–83
Bar graphs, 36–38
  vs. line graphs, 38
  stacked, 42–44
Baseline tests, 253–254
Beginning and ending of issues, 103–104
Behavioral and physiological metrics, 52, 54, 167
  Affective Computing Research Group and, 186–188
  eye-tracking, 175–180
  facial expression capture, 171–174
  overt behaviors, 167–171
  pupillary response, 180–182
  skin conductance and heart rate, 183–186
Benchmarking, 293
Benchmarking case study (mobile music and video), 263
  attempts, number of, 266
  changes and future work, 270
  comparative analysis, 264
  data collection, 265–266
  data manipulation and visualization, 267, 269

Benchmarking case study (mobile music and video) (*Continued*)
discussion, 269–270
perception metrics, 266–267
project goals and methods, 263
qualitative and quantitative data, 263
qualitative and quantitative findings, 267
research domain, 263
respondent recruiting, 265
respondents, number of, 264–265
successes and failures, 266
summary findings and SUM metrics, 267–268
time to complete, 266
Benedek, Joey, 142, 172–174
Between-subjects studies, 18–19
Biases
in issue identification, 116–117
in self-reported data collection, 126–127
"Big Opportunities" quadrant (expectation vs. experience), 131
Binary successes, 66
analyzing and presenting, 67–68
collecting and measuring, 66–67
confidence intervals, 69
Bobby accessibility checking tool, 229
Bojko, Agnieszka, 271, 279
Bots (search), 212
Brooke, John, 138
Bubb, H., 180
Budgets, 55–56, 291–292
Business goals, 295
Business language, 295
Butterfly ballots, 6, 82–83
Byrne, Michael, 148

C

Calendars, for sessions, 290
Card-sorting data, 51, 217
analysis, 218–225
closed sorts, 217, 225–228
number of participants, 223
tools, 218
CardSort tool, 218
CardZort tool, 218
Case studies, 237
metrics impact, in enterprise portal study, 280–287
mobile music and video, 263–271
pharmacist performance, 271–280
speech recognition IVR, 244–252
website redesign, 237–244, 252–253
Category, issues by, 112–113
CDC.GOV website redesign case study, 252–253
baseline test, 253–254
conclusions, 261–262
final prototype testing, 258–261
qualitative findings, 255–256
task scenarios, 255
usability testing levels, 253
wireframing and FirstClick testing, 256–258
Central tendency, measures of, 25
Chadwick-Dias, A., 201
Champions for user experience, 290
Chattratichart, J., 116–117, 119
Chi-square tests
click-through rates, 213–214
overview, 33–35
CHITEST function, 34, 214
CIF. *See* Common Industry Format
Click-through rates, 213–215
Clock operation, for time-on-task, 75
Closed card sorts, 51, 217, 225–228
Cockton, G., 119
Coding overt behaviors, 167–171
Cognitive effort, 87
Collecting data, 59–60
binary successes, 66–67
efficiency, 87–88
errors, 83–84
learnability, 93–94
levels of success, 70–72
for mobile music and video case study, 265–266
myths about, 10
self-reported data, 124–128
task success metrics, 65
time-on-task, 74–77
Color, in graphs, 35
Column graphs, 36–38
Combined metrics, 191
based on percentages, 193–198
based on target goals, 192
based on *z*-scores, 199–202
product comparisons, 50
severity ratings based on, 106–107
SUM, 202–203
usability scorecards, 203–206
Common Industry Format (CIF), 90
Comparative Usability Evaluation (CUE), 114
Comparisons
alternative designs, 55
to expert performance, 208–209
to goals, 206–208
of independent sample means, 28–29

in mobile music and video case study, 264
of more than two samples means, 30–31
of paired samples means, 29–30
products, 50
Competitor websites, 237–239
Completed transaction metrics, 48, 50
Comprehension, in self-reported metrics, 163–164
Computer System Usability Questionnaire (CSUQ), 139–140, 144–146
CONFIDENCE function, 27
Confidence intervals, 8
binary successes, 69
in culture of usability, 295–296
on graphs, 35
overview, 27–28
and sample size, 17–18, 27–28
Consistency
in data cleanup, 60–61
in identifying and prioritizing issues, 114–115
Content category, in ACSI, 151, 155
Continuous data, 22–23
Controllability area, in WAMMI, 151, 153
Converting Excel time, 75
Core measure, of efficiency, 90
Correlation function (Excel), 32
Correlations, 32–33
Costs
in culture of usability, 291–292
myths about, 10–11
COUNT function (Excel), 27
Counterbalancing, 19–20
Coyne, K. P., 108
Credible websites, 159
Criteria, for participants, 58–59
Critical product studies, 53
CSUQ. *See* Computer System Usability Questionnaire
CUE. *See* Comparative Usability Evaluation
Culture of usability metrics, 289
benchmarking for, 293–294
business language, 295
confidence intervals, 295–296
exploring data, 294
planning for, 292–293
proper use of metrics, 296–297
selling usability and metrics, 289–290
simplifying presentations, 297–298
starting small, 291
time and money, 291–292
Cynthia Says tool, 229

D

Darwin, Charles, 172
Data Analysis option (Excel), 24
Data cleanup, 60–61
Data collection. *See* Collecting data
Data Logger, 75
Data points, on line graphs, 40
Data types, 20
interval, 22–23
for metrics, 23–24
nominal, 20–21
ordinal, 21–22
ratio, 23
Decimal places, in reporting, 26
Degrees of intervalness, in self-reported data, 128
den Uyl, M. J., 172
Dependent variables, 20
Descriptive statistics, 24–25
confidence intervals, 27–28
measures of central tendency, 25
measures of variability, 26–27
Designing usability studies, 15
counterbalancing, 19–20
independent and dependent variables, 20
participant selection, 16–17
sample size, 17–18
within-subjects and between-subjects, 18–19
Desirability Rating, 174
Diamond Bullet Design case study, 231
Disabilities guidelines, 231
Display *R*-squared value on chart option (Excel), 32
Distributions, of time-on-task data, 78–79
Dixon, E., 129, 131
"Don't Touch It" quadrant (expectation vs. experience), 131
Double-counting errors, 86
Drop-off rates, 215–216
Drug label design. *See* Pharmacist performance case study
Dumas, J., 101, 114

E

Ease of use, in post-task ratings, 128–129
Effectiveness metrics, 8, 282–283
Efficiency metrics, 8, 50, 87
analyzing and presenting, 88–90
collecting and measuring, 87–88
as combination of task success and time, 90–92
frequent use of products, 51
in OneStart project case study, 282–283
in product comparisons, 50
in WAMMI, 151, 153

Effort, cognitive and physical, 87
Ekman, Paul, 171
Election ballots
  comparisons, 148–150
  errors in, 6, 82–83
Electromyogram (EMG) sensors, 172–174
Element assessment, in self-reported metrics, 161–163
EMFi chair, 185–186
Engaging websites, 159–160
Environmental bias, 116
ErgoBrowser, 75
Errors, 81
  analyzing and presenting, 84–86
  collecting, 83–84
  determining, 82–83
  issues, 86–87
  measuring, 81–84
  metric, 53
  in pharmacist performance case study, 276
  and sample size, 17–18, 27–28
Essa, I. A., 172
Evaluation methods, in studies, 57–58
Everett, Sarah, 148
Exact Method for binary success, 69
Expectation measure, in self-reported data, 129
Expectation rating, in self-reported data, 131
Experience rating, in self-reported data, 129, 131
Expert performance, comparison to, 208–209
Eye-tracking technology, 52, 175–176
  in pharmacist performance case study, 276
  scan paths, 180
  specific elements, 179–180
  specific regions, 176–179

F

F-value, in comparing means, 31
FaceReader system, 172
Facial Action Coding System (FACS), 171
Facial expression capture, 171–172
  electromyogram sensors, 172–174
  in everyday testing, 174
  video-based systems, 172–173
FACS Affect Interpretation Dictionary (FACSAID), 172
Failure scores, in binary success, 66–67
Fake issues vs. real, 101–102
Few, Stephen, 36
Filtering data, 60
FirstClick testing, 256–258
"Fix It Fast" quadrant (expectation vs. experience), 131
Fixation count and duration, in eye-tracking, 276
Florida election ballots, 82–83
Focus groups, 58
Fogg, B. J., 159
Food, at usability sessions, 290
Foresee Results, 151
Formative studies, 45–46
Formats for time data, 75
Forms, for self-reported data, 126
Frequency of issues
  unique, 109–110
  users, 111
Frequent use, of product studies, 50–51
Friesen, Wallace, 171
Frustration Index, 173–174
Fukuda, R., 180
Functionality category, in ACSI, 151, 155

G

Galvactivator glove, 185–186
Galvanic Skin Response (GSR), 183–184
Gaze plots, 176–177
GEOMEAN function (Excel), 78
Geometric mean, 78
Goals
  business, 295
  combining metrics based on, 192
  comparison to, 206–208
  in mobile music and video case study, 263
  study, 45–47
  user, 47–48
Granularity of issues, 104
Graphs
  column and bar, 36–38
  guidelines, 35
  line, 38–40
  overview, 35–36
  pie charts, 42
  scatterplots, 40–41
  stacked bar, 42–44
Greene, Kristen, 148
GSR. *See* Galvanic Skin Response
Gut level decisions, myths about, 12
Gutenberg, Johann, 7

H

Haring, S., 184
Harris, Robert, 36
Hart, Traci, 147
Harvey Balls, 204–205
Hazlett, Richard L., 172–174
Heart rate, 183–186

Heart rate variability (HRV), 183
Heat maps, of eye-tracking data, 176–177
Helpfulness area, in WAMMI, 151, 153
Hertzum, M., 108
Hierarchical cluster analysis, 220, 222
Horizontal labels, on graphs, 35
Hu, W., 183
Human Factors Research Group (HFRG), 150

I

ICA. *See* Index of Cognitive Activity
Identifying usability issues
  bias in, 117–118
  consistency in, 114–115
Imamiya, A., 183
Impact of subtle changes, evaluating, 54–55
In-person studies, 103
Independent samples, for comparing means, 28–29
Independent variables, 20
Index of Cognitive Activity (ICA), 181
Indiana University enterprise portal project case study, 280–281
  analyzing and interpreting results, 282–283
  conclusion, 287
  designing and conducting, 281–282
  findings and recommendations, 283–286
  impact, 286
Information architecture studies, 51
Information, at sessions, 290
Information Quality (InfoQual) scale (in PSSUQ), 245
Interactive Voice Response (IVR) systems
  case study. *See* Speech recognition IVR case study
  evaluations of, 149
Interface Quality (IntQual) scale (in PSSUQ), 245
International Standards Organization (ISO), 4
Interval data, 22–23
iPod, 54
Iqbal, S. T., 181–182
Issue severity, 50
Issues-based metrics, 53, 55, 99
  analyzing and reporting metrics, 108–114
  automated studies, 103
  beginning and ending of issues, 103–104
  determining issues, 100–102
  identifying issues, 99–100, 114–117
  in-person studies, 103
  multiple observers, 104–105
  number of participants, 117–121
  severity ratings, 105–108
IVR. *See* Interactive Voice Response systems

J

Jargon, 295
Jeffries, R., 101
Jenny Craig site, 88–89
Jones, Colleen, 252
Jones of New York site, 173

K

Kahneman, D., 181
Kapoor, A., 187
Kindlund, Erika, 202, 267
Krug, Steve, 4
Kuniavsky, M., 108

L

Lab tests, 57
Labeling
  axes and units, 35
  drugs. *See* Pharmacist performance case study
Landauer, T., 118
Language, of business, 295
Learnability metrics, 51, 92–93
  analyzing and presenting, 94–96
  collecting and measuring, 93–94
  issues, 96–97
  in WAMMI, 151, 153
LeDoux, L., 235
Legends, on line graphs, 40
Lekkala, Jukka, 184
Levels of success, 69–70
  analyzing and presenting, 72–73
  collecting and measuring, 70–72
Lewis, James R., 69, 129, 139, 244, 249–250, 252
Likelihood to return metric, 50
Likert scales, 124–125
Lin, T, 183
Lindgaard, Gitte, 116–117, 119, 159
Line graphs
  vs. bar graphs, 38
  guidelines, 38–40
Linkages, for card-sorting data, 220
Live-set metrics, 54
Live website data, 211
  A/B studies, 216–217
  click-through rates, 213–215
  drop-off rates, 215–216
  server logs, 211–213
  survey issues, 156–158
Look & Feel category, in ACSI, 151, 155
Loranger, Hoa, 237, 244

Lostness
in efficiency, 89–90
in information architecture, 51
Lund, Arnie, 142

M

Magic number 5, 118–120
Magnitude estimation, in self-reported data, 132–133
Management appreciation, myths about, 13
Mangan, E., 95, 235
Marsden, P., 183
Marshall, Sandra, 181
Maurer, Donna, 220
McGee, Mick, 132
McNulty, M., 201
MDS. *See* Multidimensional scaling
Means, comparing, 25, 28
independent samples, 28–29
more than two samples, 30–31
paired samples, 29–30
Means, in time-on-task, 78
Measures
of central tendency, 25
of variability, 26–27
Measuring the Usability of Multi-Media Systems (MUMMS), 151
Median
overview, 25
time-on-task, 78
MEDIAN function, 78
Meixner-Pendleton, M., 184
Memory, as metric, 52
Method selection, bias from, 116
Metrics impact case study. *See* OneStart project case study
Metrics overview, 48. *See also* Studies overview
data types for, 23–24
definition, 7–8
myths, 10–13
table, 49
value, 8–10
Miner, Trish, 142
Mobile music and video case study, 263
changes and future work, 270
comparative analysis, 264
data collection, 265–266
data manipulation and visualization, 267, 269
discussion, 269–270
number of attempts, 266
number of respondents, 264–265
perception metrics, 266–267
project goals and methods, 263
qualitative and quantitative data, 263
qualitative and quantitative findings, 267
research domain, 263
respondent recruiting, 265
successes and failures, 266
summary findings and SUM metrics, 267–268
time to complete, 266
Mode, 25
Moderator bias, 116
Molich, Rolf, 101, 114
Mota, S., 187
Mouse, pressure-sensitive, 186–188
Multidimensional scaling (MDS), 222–225
Multiple observers, benefits of, 104–105
MUMMS. *See* Measuring the Usability of Multi-Media Systems
Music. *See* Mobile music and video case study
Myths about metrics, 10–13

N

Nall, Janice R., 252, 262
NASA web site, 168
Navigation category, in ACSI, 151, 155
Navigation studies, 51
Negative relationships, 32
New products, myths about, 12
Nielsen, Jakob, 5, 106–107, 118, 213, 228
Nielsen Norman Group, 162
Noisy data, myths about, 11–12
Nominal data
nonparametric tests for, 33
overview, 20–21
Nonduplication of respondents, 158
Nonparametric tests
chi-square, 33–35
overview, 33
Nonverbal behaviors, 169, 171
Noticeability studies, 52
Null hypotheses, 29
Number
of participants, 117–121
of survey questions, 156–158
of survey respondents, 157–158
Numbered and unnumbered scales, 159

O

Objective self-reported metrics, 123
Observability, of metrics, 7–8
Observation locations, 290
Observing overt behaviors, 167–171
Omata, M., 183

OneStart project case study, 280–281
  analyzing and interpreting results, 282–283
  conclusion, 287
  designing and conducting, 281–282
  findings and recommendations, 283–286
  impact, 286
Online forms, 126
Online questionnaires, 142
Online services, 150
  ACSI, 151, 153–155
  live-site survey issues, 156
  OpinionLab, 153
  WAMMI, 150–153
Online studies, 57–58
Online surveys, 58, 126
Open card sorts, 217
Open-ended questions, 128, 163
OpinionLab, 153, 156–157
OptimalSort tool, 218
Ordering, counterbalancing for, 19–20
Ordinal data
  nonparametric tests for, 33
  overview, 21–22
Osgood, Charles E., 125
Outliers
  defined, 26
  time-on-task data, 78–79
Overloading graphs, 35
Ovo Studios, 265

P

p-values
  calculating, 120–121
  in comparing means, 29–31
Page views, 212–213
Paired samples *t*-tests, 29–30
Paired values, on scatterplots, 40
Palm Beach County, Florida, election ballots, 82–83
Paper ballots, comparisons of, 148–150
Paper forms, for self-reported data, 126
Participant knowledge, in time-on-task measurements, 80–81
Participants
  bias in identifying issues, 116
  card-sorting data, 223
  frequency of, 111–112
  number of, 117–121
  selecting, 16–17, 116
  studies, 58–59
Pentland, A. P., 172
Perceived distances, card-sorting data for, 218
Percentages
  combining metrics based on, 193–197
  in pie charts, 42
Perception metrics
  mobile music and video case study, 266–267
  self-reported data for, 123
Performance
  comparison to, 208–209
  vs. satisfaction, 48
  as user goal, 47
Performance metrics, 63–64
  efficiency, 87–92
  errors. *See* Errors
  learnability, 92–96
  pharmacists. *See* Pharmacist performance case study
  task success. *See* Successes
  time-on-task. *See* Time-on-task metrics
Perlman, Gary, 142
Pharmacist performance case study, 271–279
  analysis, 276–277
  apparatus, 272
  participants, 272
  procedure, 275
  results and discussion, 277–279
  stimuli, 272–275
Physical effort, 87
Physiological metrics. *See* Behavioral and physiological metrics
Picard, Rosalind, 184–187
Pie charts, 42
Planning studies, 45
  importance of, 293–294
  metric choices. *See* Studies overview
  study goals in, 45–47
  user goals in, 47–48
Poker, 181
Poppel, Harvey, 205
Positive findings, 114, 297
Positive user experiences, 54
Post-session ratings, 135, 137
  aggregating, 137–138
  comparison, 144–146
  CSUQ, 139–140
  product reaction cards, 142, 145
  QUIS, 139, 141
  SUS, 138–139
  Usefulness, Satisfaction, and Ease of Use Questionnaire, 142–144
Post-Study System Usability Questionnaire (PSSUQ), 139, 245–246

Post-task ratings
after-scenario questionnaire, 129–130
ease of use, 128–129
expectation measure, 129, 131
task comparisons, 133–137
usability magnitude estimation, 132–133
Posture Analysis Seat, 187
Precision
in graphs, 35
in reporting, 26
Prelaunch tests, 258–261
Presentation
binary successes, 67–68
efficiency, 88–90
errors, 84–86
learnability, 94–96
levels of success, 72–73
simplifying, 297–298
time-on-task data, 77–79
Presidential election of 2000, 6, 82–83
PressureMouse, 186–188
Printing press, 7
Prioritizing issues, 114–115
Problem discovery, 52–53, 247–248
Procedural details, 297
Product comparison studies, 50
Product reaction cards, 142, 144–145
Product use frequency studies, 50–51
Proficiency, in learnability, 93
"Promote It" quadrant (expectation vs. experience), 131
Prototype testing, 258–261
PSSUQ. *See* Post-Study System Usability Questionnaire
Pupillary response, 180–182, 276

Q

Qualitative findings
benchmarking case study, 267
CDC.GOV website redesign case study, 255–256
in mobile music and video case study, 263, 267
Quantifiability of metrics, 8
Questionnaire for User Interface Satisfaction (QUIS), 139, 141, 144–146
Questionnaires, online, 142

R

R-squared value, 32
Radar charts, 143, 204–205
Random sampling, 16
Ranges
overview, 26
time-on-task data, 78
Ranked data, 21–22
Ratings
post-session. *See* Post-session ratings
post-task. *See* Post-task ratings
for self-reported data, 127
severity, 21, 105–108
Ratios
in magnitude estimation, 132
overview, 23
Real issues vs. false, 101–102
Redesigning website case study. *See* Website redesign case study
Rehabilitation Act, 231
Relationships, between variables, 31–33
Repeated-measures studies, 18–19
Reporting. *See* Analyzing and reporting metrics
Retrospective probing technique, 80
Return on investment (ROI), 9, 231–234
Reynolds, C., 186–188
Rosenbaum, R., 95
Rubin, J., 106–107
Russell, M., 88

S

Sabadosh, Nick, 252
Safety, drug label design for. *See* Pharmacist performance case study
Safety-critical products, 196
Samples, 17–18
in comparing means, 28
myths about, 13
participants, 59
in speech recognition IVR case study, 249–250
techniques, 16
Samples of convenience, 17
Satisfaction
metrics, 8, 50
vs. performance, 48
as user goal, 47–48
Sauro, Jeff, 69, 202, 267
Scalable Vector Graphics (SVG) formats, 269
Scales
Likert, 124–125
scatterplots, 40
Scan paths, in eye-tracking, 180
Scatterplots, 32, 40–41
Scheirer, Jocelyn, 184
Schroeder, W., 119, 264
Schulman, D., 264

Scorecards, usability, 203–206
Scoring methods, in levels of success, 71
Screenshots, 297
Search and Site Performance category, in ACSI, 151, 155
Search bots, 212
Section 508 guidelines, 231
Segments, in pie charts, 42
Selecting, of participants, 16–17, 116
Self-reported metrics, 57–58, 123
  analyzing, 127–128
  attribute assessment, 158–161
  awareness, 163–165
  collecting, 124–128
  element assessment, 161–163
  online services. *See* Online services
  open-ended questions, 163
  post-session ratings. *See* Post-session ratings
  post-task ratings. *See* Post-task ratings
  rating scales for, 127
  SUS, 147–150
Self-selection, of respondents, 158
Selling usability and metrics, 289–290
Semantic differential scales, 125
Senior-friendly website comparisons, 147
Server logs, 211–213
Severity ratings, 21, 105
  caveats, 107
  combination of factors, 106–107
  user experience, 105–106
  working with, 107–108
Shaikh, A., 88
Signage disaster, 5–6
Significance level, in confidence intervals, 27–28
Simplifying presentations, 297–298
Single error opportunity, tasks with, 84–85
Single-factor ANOVA, 30–31
Single Usability Metric (SUM), 202–203
Single usability scores, 191
  based on percentages, 193–197
  based on target goals, 192
  based on *z*-scores, 199–202
  SUM, 202–203
Site map study, 161–162
Six Sigma methodology, 202, 234–236
Skin conductance, 183–186
Small improvements, myths about, 11
Small sample size, myths about, 13
Smith, Bill, 234
Smith, P. A., 89–90
Snyder, Carolyn, 116
Social desirability bias, 126–127
Software Usability Measurement Inventory (SUMI) tool, 150–151
Software Usability Research Laboratory, 147
Spearman rank correlation, 134
Speech recognition IVR case study, 244
  discussion, 251–252
  method, 244–245
  participant comments, 246–247
  PSSUQ score, 246, 282
  recommendations, 250–251
  results, 245
  sample size, 249–250
  usability problems, 247–249
Spiders (search), 212
Spool, J., 119, 264
Spreadsheet for card-sorting data, 220
Stacked bar graphs, 42–44
Standard deviation, 26–27
STANDARDIZE function (Excel), 198
Starting small, 291
Statistics, descriptive, 24–25
  confidence intervals, 27–28
  measures of central tendency, 25
  measures of variability, 26–27
STDEV function (Excel), 27, 198
Stetson, J., 144
Steven M. Ross Business School, 151
Stopping rules, for unsuccessful tasks, 73–74
Stratified sampling, 16
Stress measures, 183–186
Studies, types of, overview, 48
  alternative design comparisons, 55
  awareness, 52
  budgets and timelines, 55–56
  completing transactions, 48, 50
  critical products, 53
  data cleanup, 60–61
  data collection, 59
  evaluation methods, 57–58
  frequent use of products, 50–51
  goals, 45–47
  impact of subtle changes, 54–55
  navigation and information architecture, 51
  participants, 58–59
  positive user experience, 54
  problem discovery, 52–53
  product comparisons, 50
  table, 49
Subjective self-reported metrics, 123
Subtle changes impact, 54–55

Successes, 53, 64–65
  binary, 66–69
  data collection, 65
  in efficiency, 90–92
  issues in, 73–74
  levels, 69–73
  in time-on-task data, 80
SUM. *See* Summative Usability Metric method
SUMI. *See* Software Usability Measurement Inventory tool
Summarizing self-reported data, 128
Summary metrics, 297
Summative testing, 46–47
Summative Usability Metric (SUM) method, 202–203, 267–268
Surveys
  live-site, 156–158
  online, 58, 126
SUS. *See* System Usability Scale
SVG. *See* Scalable Vector Graphics formats
System Usability Scale (SUS), 22, 144–146
  good and bad scores, 149
  overview, 138–139
  for paper ballots comparisons, 148–150
  for senior-friendly website comparisons, 147
  for Windows ME and Windows XP comparisons, 147–148
System Usefulness (SysUse) scale (in PSSUQ), 245
Systematic sampling, 16

T

t-tests
  binary successes, 67
  mean comparisons, 28–30
Tabulating time data, 76–77
Target audience, in participant selection, 16
Target goals, combining metrics based on, 192
Task completion
  studies, 48, 50
  time. *See* Time-on-task metrics
Task-level measurements, in IVR case study, 245
Tasks
  aggregating ratings, 137–138
  issues by, 113
  with multiple error opportunities, 85–86
  selection bias, 116
  success metrics. *See* Successes
TAW Web Accessibility Test tool, 229
Technical Research Centre, 185
Tedesco, D., 133–135
Terms and jargon, 295
Text-to-speech (TTS), 250
Think-aloud protocol, 80
3D graphs, 35
Thresholds, in time-on-task data, 78
Time and time data
  collection myths, 10
  in efficiency, 90–92
  in Excel, 75
  requirements, 291–292
  savings, 233
  tabulating, 76–77
Time-on-task metrics, 74
  analyzing and presenting data, 77–79
  collecting and measuring, 74–77
  in frequent use of product studies, 51
  importance, 74
  issues, 79–81
  in pharmacist performance case study, 276
Timelines, for studies, 55–56
TiVo, 54
Top-2-box scores, 128
Top 40 Online Retail Satisfaction Index, 151
Transaction completion studies, 48, 50
Transferring data, 61
Trend lines, 32, 40
Trials, in learnability, 96–97
Trimmel, M., 184
Tufte, Edward, 36
Tullis, C., 40–41
Tullis, T.
  card-sorting data, 217, 223, 225–227
  identifying issues, 115
  learnability, 95
  self-reported metrics, 133–135, 144, 162
  Six Sigma, 235
  visual appeal ratings, 40–41
  z-scores, 201
Tylenol site, 173

U

Unique issues, frequency of, 109–110
Units, labeling, 35
Usability Information Technology Services (UITS), 280
Usability magnitude estimation, 132–133
Usability overview
  definition, 4
  importance, 5–7
  metrics, 7–8
  myths, 10–13
  value, 8–10
Usability Professionals Association (UPA), 4, 6
Usability scorecards, 203–206

Usability Testing Environment (UTE), 75, 259
Usborne, N., 216
Usefulness, Satisfaction, and Ease of Use (USE) questionnaire, 142–144
Usefulness, in self-reported metrics, 164–165
User errors metric, 53
User expectations studies, 50
User experience, severity ratings based on, 105–106
User goals, 47–48
User issues, frequency of, 111
UzCardSort tool, 218

V

Value of usability metrics, 8–10
van Kuilenburg, H., 172
Variability measures, 26–27
Variables
  creating, 60
  independent and dependent, 20
  relationship between, 31–33
Variance, 26–27
Verbal behaviors, 168–170
Verifying responses, 60
Video-based systems, for facial expression capture, 172–173
Video clips
  in presentations, 297
  for selling usability, 290
Video playback case study. *See* Mobile music and video case study
Visits, to pages, 212–213
Voice of the Customer (VoC) studies, 150

W

Wald Method, 69
Ward, R., 183
WAVE tool, 229
Web Content Accessibility Guidelines (WCAG), 229–231
WebCAT tool, 218
Website Analysis and Measurement Inventory (WAMMI), 150–153
Website redesign case study, 237
  conclusion, 244
  testing competitor websites, 237–239
  testing design concepts, 239–243
  testing single design, 243–244
Websort tool, 218
Weight-loss sites, 88–89
Weight, of graph lines, 40
Weight Watchers site, 88–89
Weighted averages, 195
Weiss, Scott, 263, 270
Whitby, Chris, 263, 270
Wilson, C., 106, 108
Windows ME and Windows XP comparison, 147–148
Wireframing, in website redesign case study, 256–258
Within-subjects studies, 18–19
Withrow, J., 232
Wolfson, Cari A., 252, 262
Wood, L., 223
Woolrych, A., 119

X

X/Y plots, 40–41

Z

z-scores
  calculating, 198
  combining metrics based on, 198–202
Zazelenchuk, Todd, 280, 287
Zheng, X. S., 181–182