Private but Misunderstood? Evidence on Measuring Intimate Partner Violence via Self-Interviewing in Rural Liberia and Malawi*

David Sungho Park Shilpa Aggarwal Dahyeon Jeong Naresh Kumar
Jonathan Robinson Alan Spearot

January 8, 2023

Abstract

Women may under-report intimate partner violence (IPV) in surveys. We conduct an experiment in rural Liberia and Malawi in which women were asked IPV questions via self-interviewing (SI) or face-to-face interviewing (FTFI). Many women appear to misunderstand questions in SI, which will likely bias reported IPV upwards (since the prevalence of a specific type of violence is typically below 50%). In Malawi, SI significantly increases reported incidence of IPV by 13 percentage points (relative to 20% in FTFI); in Liberia, reported IPV is 4 percentage points higher (on a base of 38%), but the difference is statistically insignificant.

JEL classification: C93, O12, J16, I32
Keywords: intimate partner violence (IPV), measurement, ACASI, sensitive behaviors

*Park: KDI School of Public Policy and Management (park@kdis.ac.kr); Aggarwal: Indian School of Business and J-PAL (shilpa_aggarwal@isb.edu); Jeong: World Bank (dahyeonjeong@worldbank.org); Kumar: UCSC (nkumar5@ucsc.edu); Robinson: UCSC, BREAD, CEGA, J-PAL and NBER (jmrtwo@ucsc.edu); Spearot: UCSC (aspearot@ucsc.edu). The research protocol for this study was approved by the IRBs of UCSC, the University of Liberia, and the Malawi National Committee on Research in the Social Sciences and Humanities (NCRSH). We thank USAID and IPA for funding. We are grateful to Jenny Aker for her collaboration. For organizing the data collection, we thank Joseph Davis, Arja Dayal, Wilson Dorleleay, Walker Higgins, Andreas Holzinger, Erik Jorgensen, Teresa Martens, Laura McCargo and Camelia Vasilov at IPA Liberia, and Patrick Baxter, Emanuele Clemente, Calvin Mhango, Monica Shandal, Patrick Simbewe, and Asman Suleiman at IPA Malawi. We are extremely grateful to all the enumerators in both countries, though there are too many to list individually. We thank seminar participants at Baruch College, IFPRI, UCSC, IPA-GPRL Methods and Measurement Conference 2021, and KDI School-World Bank DIME Conference 2022 for helpful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of USAID, the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1 Introduction

Intimate Partner Violence (IPV) is a pressing global public health and policy problem, but measuring its true prevalence is challenging because factors including social taboos, emotional pain, fear of retribution, or feelings of shame or embarrassment cause women to hesitate in reporting IPV to friends or family, as well as to physicians or to law enforcement officials (WHO 2012; Garcia-Moreno et al. 2013). Spurred by the lack of systematic data on IPV and recognizing its epidemiological nature, organizations such as the WHO began to run large-scale, multi-country surveys to measure the prevalence of IPV in the 1990s (WHO 1996).[1] The latest estimates from these surveys reveal that more than a quarter of ever-partnered women have experienced physical or sexual IPV during their lifetime (Sardinha et al. 2022).

Many public health professionals worry that the true rate of IPV may be higher, and that women may be understating their IPV experience even in surveys. It remains unclear if this is the case. On the one hand, some of the stigmas that drive under-reporting may be mitigated by the confidentiality afforded by a professionally done survey (as articulated in a consent form, for example), and by the fact that the surveyor is unlikely to be known by the survey respondent or her partner, or to have reason to interact with the respondent again. The survey setting also differs critically from that in normal life because the survey *directly asks* about IPV, rather than leaving the onus of initiating the conversation to the woman herself.[2] On the other hand, some of the same stigmas may still apply; for example, the victim may feel ashamed about her situation, hesitate to confide in another individual, or be scared of being overheard (despite survey precautions to guard against this).

To address some of these concerns, an alternative approach is the use of confidential self-interviewing (SI). In this approach, women self-administer IPV questions privately, which ensures that their answers are shielded even from the enumerator.[3] In this paper, we evaluate one such

[1] For example, the WHO multi-country study on women's health and domestic violence was initiated in 1997 and the DHS Program started collecting information on IPV in 1990 (the first IPV module was fielded as part of the standard DHS in Colombia).

[2] In fact, the medical literature has identified one of the key measurement approaches for IPV to simply ask the person. The WHO also recommends direct questioning as the "gold standard" method of measuring IPV. See: https://apps.who.int/iris/bitstream/handle/10665/85239/9789241564625_eng.pdf.

[3] The answers are not fully anonymized, however, since researchers have access to this data later on, but instructions during this part of the survey, and the consent form, clearly indicate that this data will be kept securely, so any risk of data breach is remote. Further, the researchers would have no reason to interact directly with the respondent outside of this research setting.

interviewing technique which is known as Audio Computer Assisted Self-Interviewing (ACASI). In ACASI, respondents listen to pre-recorded questions via headphones and respond using a touch-screen (in our setting, a tablet).[4] The enumerator has no interaction with the respondent during this part of the survey, other than to explain the module at the beginning, and to be available in case the respondent seeks clarification.[5]

The intent of ACASI is that it will destigmatize IPV reporting, which is expected to lead to an increase in reporting. However, there are two other factors which may contribute to lower IPV reporting in ACASI (especially when benchmarked against a professionally-administered survey with a trained enumerator). One, self-interviewing lacks any human element, and it is conceivable that respondents may actually be more likely to report sensitive behaviors to a human interviewer since the respondent may perceive the enumerator to be empathetic or build a rapport with her over the course of the survey.[6] If this channel is present, ACASI will actually understate IPV.

A second factor, which is the focus of our paper, is that self-interviewing requires the respondent to understand the questions on her own, and to use the tablet, which may not be easy. This is an especially salient concern in the case of IPV, as the standard set of questions for measuring IPV has fairly complex and nuanced language, and therefore it may not be straightforward to grasp without the surveyor helping with interpretation. In almost every setting, misunderstanding will tend to cause IPV to be *over*-reported. This is because IPV is measured through a module containing 20 questions which are later indexed into 4 main categories (controlling behavior, emotional violence, physical violence, and sexual violence). Typically, the mean of each of these individual yes/no questions is well under 0.5, so a woman who does not understand the module and randomly answers yes or no will tend to bias the level of reported IPV on any given question *upwards*. This bias will be amplified in the indexing (which is set equal to 1 if the respondent reported any form of IPV).

To shed light on these various channels, we conduct a measurement experiment within surveys collected as part of an evaluation of an unconditional cash transfer program in rural Liberia and

[4] Another reason to recommend ACASI is that it is virtually impossible for the interview to be overheard by anyone in close vicinity. However, this is not relevant in our experiment (or in any survey which uses best practice face-to-face interviewing), since the survey is always conducted privately.

[5] In our survey protocol, the respondent could pause the module to ask questions, and the enumerator could help her to resume from where she left off.

[6] Indeed, Ellsberg et al. (2001) compile anecdotes from debriefings of IPV survey enumerators in Nicaragua recounting how they were moved or distressed by the respondents' IPV experiences, and some even reported that respondents sought their counsel during or at the end of the IPV module.

Malawi (Aggarwal et al. 2023). Women were individually randomized into whether the IPV module was asked either via face-to-face interviewing (FTFI) or over self-interviewing (SI). Baseline IPV rates differ dramatically across the two samples: the proportion of women experiencing any type of IPV over the past year (measured in FTFI) is 20% in Malawi but 38% in Liberia; as such, we opt to present all results separately by country.

We have three main findings. First, we check for respondents' understanding of the SI tool through 5 non-sensitive screening questions, for which the answer should universally be yes. These were administered to *all* respondents through SI, irrespective of the modality through which they were asked the IPV module. The specific wording of these questions was developed over pre-testing and was meant to be understandable among the study populations. These questions are (1) Are you a woman? (2) Do you live in the [location where the survey is being conducted]? (3) In the past week, did you sleep, during day or night? (4) In the past year, did it rain in your village one time or more? and (5) Have you heard about the Coronavirus? We find that a sizeable fraction of women incorrectly answer "no" to these questions, with lower rates of errors for the more obvious questions such as gender or location and rates as high as 14-22% for the rain and sleep questions.[7] Even the responses to the most basic questions on gender and location are not unanimously affirmative, with 2-5% of the women making errors on the gender question and nearly 10% doing so on location. Overall, only 62-70% answer all 5 questions correctly, and only 84-88% answer the 3 simplest questions correctly (gender, rain, and knowledge of the coronavirus).[8]

Second, after screening, further questions were randomized to be administered either by FTFI or SI. As part of this module, we included a further set of innocuous "placebo" questions; since these are administered either by FTFI or SI, we can estimate placebo treatment effects. The placebo includes 4 questions for which the answer could be yes or no: (1) Did you do any farm work in the past year? (2) Did you go to the market in the past week? (3) Will you, or anyone in your household, eat any [rice/maize] next week, one time or more? and (4) Will you, or anyone in your household, eat any type of meat next week, one time or more? If SI is accurate, we should find

[7] The research team found such responses even during pre-testing, and so repeatedly refined the surveys, including iterating on the exact language of these questions; however, such reporting remained. While we have no definitive answer to why women answer as they do, one anecdote is that women with young children or who are nursing interpret "sleep" as being about getting restful sleep. We have no explanation for the rain question.

[8] In phone surveys that we ran with a random subset of these respondents in the immediate aftermath of the COVID-19 lockdowns (on topics unrelated to IPV), we found universal awareness about the virus and specifically, that it was called the coronavirus (Aggarwal et al. 2022).

no SI effects on these placebo questions, at least for those who understood the screening questions mentioned above. To examine this, we conservatively define women as having "passed" screening if they answered correctly the 3 most widely understood questions (gender, location, and having heard of the Coronavirus). Yet, we find placebo effects even for those who passed screening: while the size of the placebo effect is smaller in magnitude for those who passed, it is consistently statistically significant. In this context, the screening questions were evidently not extensive enough to predict comprehension, a situation no different than standard SI protocols, which typically do not check for comprehension.

Third, we find that SI increases IPV reporting, but that this increase may be entirely spurious. The increase is quite sizeable, for all categories of IPV (i.e. controlling behavior, emotional, physical, and sexual IPV): on a given question, 7% of women in Malawi and 14% in Liberia report yes in FTFI, and SI increases this percentage by 5 percentage points in Malawi and 3 percentage points in Liberia. As an index, the effects are even larger, at least in Malawi, where the probability of emotional, physical or sexual IPV increased by 5-10% points, on a base of 7-16%. In Liberia, the effects are more modest: 1-8% points (significant only for sexual IPV) on a base of 7-34%. Naively interpreted, the increase in IPV we document would match the narrative that women are hesitant to report IPV, and that FTFI dramatically understates prevalence. However, we know from the screening questions that a significant portion of women do not seem to understand the ACASI module, and the effect sizes for our placebo results are similar to those for IPV. Our interpretation is that ACASI is not appropriate, at least for these populations, and researchers should be extremely cautious about using ACASI.

Our paper is related to a large but as yet inconclusive literature about the effects of ACASI on measuring sensitive behaviors. Studies comparing ACASI and FTFI in a variety of contexts suggest that self administration increases the reporting of sensitive behaviors.[9] However, since researchers typically do not have an objective measure of the underlying behavior, it is not clear whether this increase is indicative of increased truthful reporting, or miscomprehension. Falb et al. (2016) tested ACASI with adolescent girls in the Democratic Republic of Congo (DRC) and

[9]See Tourangeau and Yan (2007) for a review. There are also examples where FTFI is more effective in some contexts. For example, Fincher et al. (2015) find evidence that FTFI is more effective than ACASI in screening women for IPV in WIC clinics in the US. In a study set in the context of a syringe exchange program, Newman et al. (2002) find that ACASI increases reporting of stigmatized behaviors but decreases reporting of psychological distress.

in refugee camps along the Sudan-Ethiopia border, and report that self-reported *average* ACASI comprehension levels are only 90% for the DRC and 75% for the Sudan-Ethiopia border, a level similar to our study. Park and Kumar (2022), a concurrent study to ours in Monrovia, Liberia, find that even urban and educated women struggle with ACASI comprehension.

Our paper is also related to a broader recent literature about survey methodologies aimed at preserving respondent confidentiality; ACASI is only one of these. Other methods rely on indirect responses, such as list experiments or randomized response techniques.[10] There is no consensus on the efficacy of these methods (see Höglinger and Jann 2018 and Lensvelt-Mulders et al. 2005 for reviews). For example, Chuang et al. (2021) find logical errors in a list experiment and a randomized response technique focused on sexual and reproductive behavior. Agüero and Frisancho (2021) finds no difference in IPV reporting between the two methods among urban microfinance borrowers in Lima, Peru. Cullen (2020) randomizes FTFI, list randomization, and ACASI in Rwanda, and find some evidence that ACASI increases reported violence modestly (and that list randomization increases it much more). Researchers have also tried other unconventional methods to measure IPV indirectly, such as asking female community leaders, but these efforts have not been very successful (Agüero et al. 2020).

The rest of this paper is laid out as follows. Section 2 describes the experiment and data collection. Section 3 presents our main results. Section 4 discusses evidence on potential pathways and heterogeneity. Section 5 concludes.

2 Data and Experimental Design

2.1 Setting

The ACASI experiment we analyze was done as part of an endline survey for a cash transfer RCT in Liberia and Malawi (the transfers were implemented by the NGO GiveDirectly, as part of a USAID-funded study, Aggarwal et al. 2023). The study takes place in Bong and Nimba counties

[10] In list experiments, yes/no questions about sensitive behaviors are included in a list with other innocuous binary-response questions, and subjects report the number of items for which the answer is "yes" or "no", which allows the researcher to back out the population level prevalence of a behavior without being able to identify whether a specific individual engaged in that behavior. The randomized response technique (RRT) bundles a question with a random event, such as a throw of the dice. Respondents are instructed to report truthfully only if the die landed on a certain number.

in Liberia, and in Chiradzulu and Machinga districts in Malawi. The study includes 300 villages in each country, with half of the villages receiving cash transfers worth $500 on average. While we do not evaluate the transfers themselves in this study, one important detail is that villages were included in the study only if they fell below a population threshold, as measured in the most recent population census (in Malawi, the upper threshold was 100 households per village; in Liberia, it was 125). The reason for this is that transfers were given out universally in treatment villages, and so our partner NGO chose smaller villages to be able to preserve their liquidity.

In Liberia, we implemented the project in two waves: a smaller first wave (90 villages), which had its endline in late 2020; and a bigger second wave (210 villages), which had its endline in September-November 2021. Most of our ACASI protocols were developed, tested, and refined over the course of the Wave 1 endline. Therefore, this sample is excluded from our results, and our results for Liberia are restricted to Wave 2 only. In Malawi, all 300 villages were enrolled at once and the endline was in April-July 2021. Figure A1 presents the project timeline.

In both countries, we attempted to enroll 10 households per village into data collection for program evaluation, though in some cases we were only able to enroll fewer households. Surveys were targeted at female heads of households, i.e., either the spouse or partner of the male head for dual-headed households or in rare cases, the solo head for female-headed households (as we show in Table 1, 97% of the sample in both countries had a partner at the time of our study). Male heads of dual-headed households were interviewed only when the female was not present, and could not be reached within a few days of the initial visit; when the male head was interviewed, the IPV module was not asked.

2.2 Questionnaire Design and ACASI Experiment

Measuring Intimate Partner Violence

To measure IPV, we employed WHO's standard Violence Against Women module.[11] The questionnaire includes 20 questions about experience with specific forms of violence, over a time period of 12 months prior to the survey. Following the literature, we group these questions into four categories:

[11] The WHO's standard questionnaire for measuring IPV, which is widely used for measuring IPV, can be found here: https://www.who.int/gender/violence/who_multicountry_study/Annex3-Annex4.pdf. Our survey module on IPV can be found in Appendix F).

controlling behavior, emotional IPV, physical IPV, and sexual IPV. In conducting this module, we followed WHO's ethics protocol for IPV research, which includes hiring only female enumerators; training enumerators to safely conduct the interviews and to be prepared emotionally for the work; conducting all surveys privately; reiterating consent just before the IPV module; and providing all respondents with an information sheet that listed the services available for women experiencing IPV, including law enforcement and local hospitals (WHO 2016).

The IPV module was attempted to be administered to all women who had an intimate partner within the 12 months preceding the survey. A small percentage (less than 1%) of women refused the IPV questions entirely, however. In addition, for the ACASI experiment, we excluded all women who reported having vision or hearing impairment (since they could not take the ACASI module); these women were administered the module via FTFI and are not included in this paper. About 7% of our sample were excluded for this reason.[12] We also excluded the Liberia Wave 1 sample because most of the ACASI survey protocols were developed and refined throughout that period. With all these restrictions, the sample size for this paper is 2,998 women (1,737 in Malawi and 1,261 in Liberia).

ACASI Implementation

In ACASI, respondents listen to questions on headphones and answer questions privately on a tablet. In each country, audio readings of the questions were recorded by an enumerator who was chosen for having clear enunciation. The recorded audio files were uploaded to SurveyCTO, along with image files for choice options (i.e. "yes" / "no" / "refuse to answer" / "don't know"). As shown in Figure A2, the resulting interface on the tablet has a speaker icon (which the respondent could touch to listen to the question) and four images (from which the respondent could choose her answer by touching the screen herself).

In the field, the enumerator explained how to take the module, and then demonstrated how to conduct the module by going through a handful of practice and demonstration questions with the respondent, and making sure that she could clearly hear the audio and accurately choose the option she intends to. When the respondent felt ready to take the actual module, the enumerator handed the tablet over to the respondent for her to take the module. In order to make sure that she had

[12] In Table A1, we show that the likelihood of these exclusions was balanced between treatment and control.

complete privacy while doing so, the enumerator kept sufficient distance to be unable to see the screen but remained in the same room or vicinity to be available to answer questions. When the respondent handed back the tablet, the screen was blank so that the responses were blinded to the enumerator.

Experimental Design

In each survey round for each country, half of the sample was randomly assigned to ACASI and the other half to FTFI. However, before starting the IPV questions, *every* respondent was asked to take a set of "screening" questions via ACASI.[13] The answers to all of these questions are expected to be yes: (1) "Are you a woman?" (2) "Do you live in [the county/district in which the survey is being conducted]?" (3) "In the past week, did you sleep, during day or night?" (4) "In the past year, did it rain in your village one time or more?" and (5) "Have you heard about the coronavirus?"[14]

After screening, women began questions in their experimental group (either ACASI or FTFI). As discussed throughout the paper, this module included questions on IPV, but it also included "placebo" questions. These placebo questions were meant to be innocuous and free from any stigma or social desirability bias, and to be a further tool to calibrate the effects of ACASI. They included 4 questions: (1) "Did you do any farm work in the past year?" (2) "Did you go to the market in the past week?" (3) "Will you, or anyone in your household, eat rice/maize in the next week, one time or more?"[15] and (4) "Will you, or anyone in your household, eat meat in the next week, one time or more?" Though the wording of these questions may look cumbersome in English, the specific language was developed after pre-testing to best be understood by the study populations.

Other Subtreatments

To explore possible technical reasons for misunderstanding, we cross-cut multiple sub-treatments. First, we randomized whether the "yes" or "no" option would appear at the top of the screen (Figure A3). This randomization was implemented in order to test whether respondents are more

[13] These screening questions were added after piloting, when it became apparent that women were answering unexpectedly to innocuous placebo questions.

[14] We also asked one question that would likely be answered "no": "Have you traveled outside the country in the past week?" We do not use this in our main specifications, however, because some women could potentially travel across borders (especially in Nimba county in Liberia, which borders Guinea, and Machinga district in Malawi, which borders Mozambique), due to which "no" is not a perfect benchmark.

[15] This question is about the staple food, which is maize in Malawi and rice in Liberia.

or less likely to pick the first option. Second, in order to examine possible learning effects in which respondents became more comfortable with the method with more experience, we randomized whether the placebo questions come before or after the IPV module.[16]

2.3 Summary Statistics and Balance Check

Table 1 shows summary statistics by country sample, as well as the difference between the SI and FTFI groups. Panel A shows household demographics. Because the sample is restricted to women with an intimate partner at any point during the past 12 months, the proportion of women who are currently partnered is very high (97%). The average respondent is about 37-38 years old and lives in a household with 5-6 members. Panel B shows education and mobile phone ownership. Average educational attainment is 5.2 years in Malawi and only 2.4 in Liberia. Sixty-six percent of women in Malawi are literate, compared to 30% in Liberia. Mobile phone ownership is similar in the two countries, ranging from 42% in Liberia to 45% in Malawi.

Panel C shows some indicators of household income and wealth, and reveals that households are better off in Liberia than in Malawi: average total monthly household expenditures are $26 in Malawi and $66 in Liberia, or about $0.17-0.39 in per capita daily expenditures. In Malawi, the average household reports about $160 worth of assets, compared to $420 in Liberia. Most of the households in the study villages are subsistence farmers, and the average monthly non-agricultural income measured in our surveys is $8-10.

Panel D shows a few proximate indicators related to female empowerment. Forty-four percent of women in Malawi have their own income source, compared to 31% in Liberia. The age difference (in years) between husband and wife is 2.9 in Malawi and 4.1 in Liberia.

Turning to Columns 2 and 4, we find two outcomes for which the differences are significant at 10% in Malawi (food security and total expenditure, which are both lower in the ACASI group), and none in Liberia. While the randomization appears to show no cause for concern, we present results separately with and without controls, and find no significant difference in results. In any case, we control for all variables reported in Table 1 for the main analysis.

[16] While we ended up not including this in the analysis for this paper, we also randomized the order between the IPV questionnaire and another likely sensitive module, the Patient Health Questionnaire (PHQ-9), which measures psychological well-being. We control for this cross-randomization in relevant analysis shown in Appendix C.

3 Results

3.1 Screening Questions

We start by documenting responses to the five screening questions which were administered to all respondents via ACASI. Results, which are shown in Table 2, suggest major cause for concern. Only 95-98% report being a woman, and 91-93% report living in their county/district of residence. Even more surprisingly, only 78-86% report that they slept in the past week and 83-85% report that it rained in the past year. While we do not have a good explanation for these results, an *ex post* explanation from some of our field staff was that some women interpreted the sleep question as "getting a good night's sleep," to which some women reported no. A parallel *ex post* explanation for the rain question is that women may have interpreted it as meaning whether it rained "enough." The reasons for misinterpretation of these questions notwithstanding, the bottom line is that even these *simple* questions were very likely misinterpreted, raising concerns about how well the more nuanced IPV questions would be understood.

Taking the questions together, we find that only 62% of respondents in Malawi and 70% in Liberia correctly answered all the questions. If we restrict to the 3 simplest questions (gender, location and COVID), these numbers become 84% and 88%. Either way, these responses suggest that many women will not be able to use ACASI effectively, and that it will be difficult to estimate a population level prevalence using ACASI. As such, we could also not definitively identify sub-populations where ACASI will work well, at least in our sample - we show correlations between "passing" screening (defined as answering the 3 simpler questions) and respondent characteristics in Table A2. We see that in Malawi less educated women are more likely to "fail" the screening (though in Liberia the correlation is insignificant).

3.2 Placebo Effects

In this section, we examine effects of SI on placebo questions, and test whether we can use the screening questions to predict which women are more likely to answer correctly using SI. To do this, for each country sample, we run the following regression:

$$Y_i = \beta SI_i \times ScreenPass_i + \gamma SI_i \times NonPass_i + \delta ScreenPass_i + \boldsymbol{X}_i'\boldsymbol{\theta} + \varepsilon_i, \tag{1}$$

where $ScreenPass_i$ is equal to 1 if individual i chose yes to all of the 3 simplest questions in Table 2 (gender, location, and coronavirus), and 0 otherwise (and $NonPass_i$ is the complement). $\boldsymbol{X}$ is a vector of covariates including all variables in Table 1.[17] The coefficients of interest are β, which represents the SI effect for those who passed the screening, and γ, which is the SI effect for those who did not. We also present p-values for a test of equality of β and γ.

Results are presented in Table 3, separately for Malawi (Panel A) and Liberia (Panel B).[18] In either country, we cannot reject equality of effect sizes for those that passed and those that didn't, for any question. At the bottom of each panel, we show the effect of SI for the average respondent (i.e. a weighted average of β and γ), and test for significance. In Malawi, effects are highly significant for 3 of 4 outcomes, and economically large. The pooled SI effects are significant for 2 out of 4 outcomes in Liberia.

These results are conditional on answering yes or no; we drop those that report "don't know" or who refused to answer. In Table A3, we examine whether the probability of responding yes or no is correlated with ACASI. We find a small increase in responding don't know or refuse to answer for farm work and market visit in Malawi, but a decline in eating maize or rice. In Liberia, we only see a relatively small effect on eating rice. We prefer to show the main results conditionally, however, so that they are more easily interpretable.[19]

In Table A5 and Table A6, we examine heterogeneity in placebo effects with background characteristics that might be correlated with being able to complete the module, including education, mobile phone access, literacy, and age. Of the 32 regressions across both tables, only four have a p-value for equality of estimated treatment effects below 0.1 (and only one below 0.05). Even qualitatively, we see no pattern in Malawi. In Liberia, we find some suggestive evidence that ACASI might be more effective with educated younger women (similar to the finding in Falb et al. 2016). However, given the contrast in findings from Liberia and Malawi, taken together, these results again underscore the difficulty in making any *ex ante* judgments about the suitability of ACASI for a given context or population.

In Table A7 and Table A8 we examine other correlates which might be related to attention

[17] Results without controls are shown in Appendix D, which are essentially identical.

[18] Table A4 runs a version of these results with the additional screening question on travel outside the country. Results are very similar with this screening definition.

[19] In Appendix E, we show "unconditional" results, imputing "don't know" and "refusal" responses with zeros instead of dropping them. These are virtually identical to the main results.

span, notably whether the survey was in the morning or afternoon (Panel A) and food security (Panel B). While a few coefficients are significant, the pattern is inconsistent across countries and across outcomes.

3.3 Effect of ACASI on IPV Reporting

Next, we show the ACASI effects on the main outcome of interest, IPV. Because we find placebo effects even for those who passed screening, we pool all women in this analysis.[20] We first estimate a regression at the *question* level:

$$IPV_{iq} = \beta SI_i + \boldsymbol{X}_i'\boldsymbol{\theta} + \psi_q + \varepsilon_{iq} \tag{2}$$

where IPV_{iq} is the binary indicator of whether individual i responded yes to question q, and ψ_q question-level fixed effects. All other notation is the same as Equation (1). We report results separately for each category of IPV: controlling behavior, emotional IPV, sexual IPV, and physical IPV. In a second analysis, we estimate the same equation but for the IPV *index*, which is set equal to 1 if a respondent reported violence on *any* question in that category.

The question-level results are presented in Table 4. For Malawi (Panel A), all effects are statistically significant, and range between 1 (physical IPV) and 9 percentage points (controlling behavior). A specification that pools all question categories together (Column 5) finds a 5 percentage point increase in reporting, also significant. In Liberia (Panel B), effects are slightly more modest, where only 3 of 5 coefficients are significant, and effect sizes range from 1-5 percentage points. However, these effect sizes are similar to those for the placebo effects, and therefore, we cannot rule out the possibility that they may be driven by miscomprehension of the module.

In Table 5, we show results at the index level. The findings are qualitatively similar to the ones for individual questions, although results here differ dramatically by country. In Malawi, ACASI increases emotional IPV by 10 percentage points (base 16%), physical IPV by 5 percentage points (base 8%), and sexual IPV by 6 percentage points (base 7%). Across all forms of IPV (not including controlling behavior), ACASI increases prevalence by 13 percentage points, a 65% increase on the base of 20%. In Liberia, effects are positive but surprisingly much more modest: the index of any

[20] Disaggregated results are shown in Table A9. However, we find little meaningful variation.

form of IPV increases by only 4 percentage points (on the much higher base of 38%).[21]

To sum up our findings, we find clear evidence that SI dramatically increases IPV reporting (at least in one country, Malawi). While it is possible that some of this increase is indeed indicative of destigmatization, it is also entirely possible that the effects are driven purely by comprehension difficulties. Our results suggest that caution is warranted in using ACASI, at least in settings like these.

4 Investigation of Heterogeneity and Pathways

Debriefing: Did Technical Problems Impede Understanding?

A simple hypothesis for these results is that technical problems made it hard to understand or complete the ACASI module, and therefore, a technically superior module may eliminate the purported miscomprehension. We believe that this is unlikely as before implementing these protocols, we extensively pre-tested the modules, especially after early results showed similar patterns to those reported here. We carefully tested that the audio instructions were well articulated and read at a reasonable speed, and refined the implementation over time. Nevertheless, technical difficulties could have remained.

To evaluate this, after the respondent handed back the tablet to the enumerator, she asked a handful of debriefing questions about whether the respondent had faced any technical or comprehension difficulties during the module, which we present in Appendix B. As shown in Table B1, only 1-2% reported technical issues; most respondents could hear the module, and felt the recordings were slow enough to understand. In Table B2, we regress answers to these technical questions on passing screening. We find no correlation here, which is perhaps not surprising given the low level of technical difficulties. We find no evidence that simple technical problems were the explanation.

On the other hand, we show in Table B3 that 8-12% reported comprehension difficulties with the module, in remembering which picture meant "yes" (a green check) and which meant "no" (a red cross), or in using the tablet. In Table B4, we regress passing screening on these measures of self-reported comprehension. In both countries, we see that people who reported understanding

[21] In Table A10, we analyze whether ACASI increases the likelihood of respondents picking "don't know" or "refuse to answer." We find that the probability of these answers is miniscule in FTFI, and SI leads to a small increase in these being chosen. Moreover, as is evident from footnote 19, our results are not driven by this selection/attrition.

the module were more likely to pass screening (though significantly so only for one measure). This is consistent with the idea that some people had trouble understanding how the module worked.

Subtreatments

While informative, these are only debriefing questions. To shed further light on this, we randomized several subtreatments to evaluate technical components of the module (results are presented in Appendix C). First, to examine whether the location of the choice options on the screen affects reporting, we randomized the order of the yes and no options. This subtreatment was motivated by our suspicion that when in doubt, some women may have the tendency to simply choose the first option. We start by analyzing this for the placebo questions in Table C1, and find evidence that respondents were more likely to choose yes when it appears at the top of the choice options in Malawi for 2 of 4 placebo questions, although not in Liberia. Surprisingly, however, in Table C2, we find no evidence of the presence of such behavior in either country when it comes to the IPV questions. We have no good explanation for why this may be the case.

Second, in order to check for the possibility that respondents may get better at understanding the module with practice, we randomized the order between the non-sensitive placebo questions and the IPV questions. Specifically, for half the sample (cross-randomized into FTFI and SI), the IPV questions came before the placebo ones, while for the other half, this order was reversed. For the placebo questions (Table C3), we find no effect of ordering, other than for the farm work question in Malawi. However, the effect goes contrary to the expected direction as the placebo effect of SI comes about when the placebos come later (i.e., practice does not help). That said, we do not wish to make much of this lone coefficient, as the placebo effect of SI is not significantly different between "placebos first" and "IPV first" in all other cases. For the IPV questions, we report coefficients in Table C4, and find that IPV reporting increases for sexual IPV in Malawi if the placebos come first. Overall, for Malawi, the effect of SI on the probability of answering a question "yes" is about 4 percentage points if the IPV questions come first, but 6 percentage points if the placebos come first (p-value for difference = 0.217). We find no significant effect of the ordering in Liberia either. This finding is consistent with the possibility that survey fatigue causes measurement error to increase, though it is also possible that the increase in IPV is real and that women became more familiar with the module over the course of the survey. We leave a further investigation of this channel to

future work.

5 Conclusion

In this paper, we test the efficacy of SI versus FTFI in eliciting truthful responses regarding IPV from women respondents in the context of a cash transfer experiment in rural Liberia and Malawi. Our results suggest that women do not understand SI, as evidenced by the fact that a substantial fraction of women answer basic placebo questions incorrectly. This lack of understanding will tend to *increase* IPV reporting, since the rate of IPV on an individual question is typically much less than 50%. And indeed, we do find a striking increase in reported IPV in one country (Malawi). However, this result is likely entirely spurious. This is deeply concerning because measurement error goes in the same direction as destigmatization, and so what looks like a decrease in stigma could be purely fictional. SI could therefore give very misleading results.

Our results, combined with our read of the literature, suggest that there may be greater benefit from having well-trained, empathetic enumerators than from SI in the context of measuring IPV. For example, in a natural experiment in Serbia, respondents of a WHO-run IPV survey ended up getting randomly assigned to either a previously inexperienced but well-trained enumerator (training duration of 2.5 weeks) or to an experienced, professional enumerator, but with less than a day of IPV training.[22] While 21% of the women reported having experienced physical or sexual IPV to the untrained enumerators, 26% reported IPV to the trained ones (Jansen et al. 2004).

Another relevant data quality issue from a companion study in the same setting (Jeong et al. 2023) is that time into the survey at the point at which a question is asked appears to adversely impact response quality, a phenomenon known as survey fatigue. While fatigue would be a consideration for any survey, it may be particularly germane for IPV measurement because many surveys choose to place the most sensitive modules at the end of a survey. For example, the standard DHS surveys ask about domestic violence at the end; we also chose to always place the IPV module at the end, even as we randomized the location of other survey modules within the survey. While this is usually done to minimize shame or embarrassment stemming from continued interaction with the enumerator after having answered the IPV module, it may unintentionally worsen data quality.

[22] This was done in an effort to speed up the fieldwork midway through surveying after the assassination of then Prime Minister Zoran Đinđić in March 2003.

In sum, therefore, non-conventional methods to collect data about stigmatized behaviors should be implemented with caution as they may open up unexpected channels of bias. It is also advisable to accompany new methods with extensive testing and other ways of ground-truthing prior to widespread implementation.

References

Aggarwal, Shilpa, Jenny Aker, Dahyeon Jeong, Naresh Kumar, David Sungho Park, Jonathan Robinson, and Alan Spearot (2023). "The Dynamic Effects of Cash Transfers: Evidence from Rural Liberia and Malawi." Unpublished.

Aggarwal, Shilpa, Dahyeon Jeong, Naresh Kumar, David Sungho Park, Jonathan Robinson, and Alan Spearot (2022). "COVID-19 market disruptions and food security: Evidence from households in rural Liberia and Malawi." *PLOS ONE* 17 (8): 1–19.

Agüero, Jorge M., Úrsula Aldana, Erica Field, Veronica Frisancho, and Javier Romero (2020). "Is Community-Based Targeting Effective in Identifying Intimate Partner Violence?" *AEA Papers and Proceedings* 110: 605–609.

Agüero, Jorge M. and Veronicar Frisancho (2021). "Measuring Violence Against Women with Experimental Methods." *Economic Development and Cultural Change*.

Chuang, Erica, Pascaline Dupas, Elise Huillery, and Juliette Seban (2021). "Sex, lies, and measurement: Consistency tests for indirect response survey methods." *Journal of Development Economics* 148: 102582.

Cullen, Claire (2020). "Method Matters: Underreporting of Intimate Partner Violence in Nigeria and Rwanda." World Bank Policy Research Working Paper 9274.

Ellsberg, Mary, Lori Heise, Rodolfo Pena, Sonia Agurto, and Anna Winkvist (2001). "Researching Domestic Violence against Women: Methodological and Ethical Considerations." *Studies in Family Planning* 32 (1): 1–16.

Falb, Kathryn et al. (2016). "Implementation of Audio-Computer Assisted Self-Interview (ACASI) among adolescent girls in humanitarian settings: feasibility, acceptability, and lessons learned." *Conflict and Health* 10 (32).

Fincher, Danielle, Kristin VanderEnde, Kia Colbert, Debra Houry, L. Shakiyla Smith, and Kathryn M. Yount (2015). "Effect of Face-to-Face Interview Versus Computer-Assisted Self-Interview on Disclosure of Intimate Partner Violence Among African American Women in WIC Clinics." *Journal of Interpersonal Violence* 30 (5): 818–838.

Garcia-Moreno, Claudia, Christina Pallitto, Karen Devries, Heidi Stockl, Charlotte Watts, and Naeema Abrahams (2013). *Global and Regional Estimates of Violence Against Women: Prevalence and Health Effects of Intimate Partner Violence and Non-partner Sexual Violence*. World Health Organization.

Höglinger, Marc and Ben Jann (2018). "More is not always better: An experimental individual-level validation of the randomized response technique and the crosswise model." *PLOS ONE* 13 (8).

Jansen, Henrica, Charlotte Watts, Mary Ellsberg, Lori Heise, and Claudia Garcia-Moreno (2004). "Interviewer Training in the WHO Multi-Country Study on Women's Health and Domestic Violence." *Violence against Women* 10 (7): 831–849.

Jeong, Dahyeon, Shilpa Aggarwal, Jonathan Robinson, Naresh Kumar, Alan Spearot, and David Sungho Park (2023). "Exhaustive or exhausting? Evidence on respondent fatigue in long surveys." *Journal of Development Economics* 161: 102992.

Lensvelt-Mulders, Gerty J. L. M., Joop J. Hox, Peter G. M. van der Heijden, and Cora J. M. Maas (2005). "Meta-Analysis of Randomized Response Research: Thirty-Five Years of Validation." *Sociological Methods & Research* 33 (3): 319–348.

Newman, Jessica Clark, Don C. Des Jarlais, Charles F. Turner, Jay Gribble, Phillip Cooley, and Denise Paone (2002). "The Differential Effects of Face-to-Face and Computer Interview Modes." *American Journal of Public Health* 92 (2): 294–297.

Park, David S. and Naresh Kumar (2022). "Reducing Intimate Partner Violence: Evidence from a Multifaceted Female Empowerment Program in Urban Liberia." Working Paper.

Sardinha, Lynnmarie, Mathieu Maheu-Giroux, Heidi Stöckl, Sarah Rachel Meyer, and Claudia García-Moreno (2022). "Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018." *Lancet (London, England)* 399 (10327): 803–813.

Tourangeau, Roger and Ting Yan (2007). "Sensitive questions in surveys." *Psychological Bulletin* 133 (5): 859–883.

WHO (1996). "Violence against women : WHO consultation, Geneva, 5-7 February, 1996."

— (2012). "Understanding and addressing violence against women: Intimate partner violence." World Health Organization Report.

— (2016). "Ethical and safety recommendations for intervention research on violence against women." World Health Organization Report.

Table 1: Summary Statistics and Experimental Balance

	(1)	(2)	(3)	(4)
	Malawi[a]		Liberia[b]	
	FTFI	SI	FTFI	SI
	Mean [SD]	- FTFI	Mean [SD]	- FTFI
Panel A. Demographics				
=1 if currently married or has partner	0.97	-0.01	0.97	-0.00
		(0.01)		(0.01)
Age	37.97	-0.94	37.13	0.67
	[12.88]	(0.60)	[10.96]	(0.61)
Number of household members	5.03	-0.02	5.59	-0.09
	[1.78]	(0.09)	[2.27]	(0.13)
Panel B. Education and mobile phone ownership				
Years of education	5.22	0.01	2.44	-0.01
	[3.50]	(0.17)	[3.43]	(0.19)
=1 if able to write/read	0.66	-0.01	0.30	0.01
		(0.02)		(0.03)
=1 if has access to mobile phone	0.45	0.03	0.42	-0.01
		(0.02)		(0.03)
Panel C. Household wealth				
Food security index (z-score)	0.00	-0.09*	0.00	0.05
	[1.00]	(0.05)	[1.00]	(0.06)
Total expenditure (monthly)	26.03	-2.13*	65.71	-0.52
	[24.46]	(1.17)	[47.08]	(2.59)
Net value of durables, livestock, and financial asset	162.55	4.24	416.43	33.88
	[235.93]	(11.45)	[823.80]	(51.15)
Non-agricultural income (monthly)	10.27	0.96	7.84	0.85
	[16.73]	(0.81)	[20.52]	(1.13)
Panel D. Empowerment-related outcomes				
=1 if has her own income source	0.44	0.01	0.31	-0.03
		(0.02)		(0.03)
Age difference from spouse	2.94	-0.16	4.09	0.72
	[10.78]	(0.51)	[12.59]	(0.72)
Observations	1,737		1,261	

Note: Sample is restricted to women with an intimate partner over the 12 months prior to the survey, and those who do not report any vision or hearing impairments. Columns 1 and 3 present the mean for the FTFI groups, and Columns 2-4 show the difference between the ACASI and FTFI groups. Standard deviation is in square brackets in Columns 1 and 3 and standard error in parentheses in Columns 2 and 4. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table 2: Self-interviewing (SI) Screening Questions

	(1)	(2)
	Mean (=1 if yes)	
	Malawi	Liberia
Questions for which answer should be yes:		
1. Are you a woman?	0.95	0.98
2. Do you live in [the county/district where the survey is being conducted]?	0.91	0.93
3. In the past week, did you sleep, during day or night?	0.78	0.86
4. In the past year, did it rain in your village one time or more?	0.83	0.85
5. Have you heard about Coronavirus?	0.93	0.94
Summary measures for "passing" screening		
=1 if YES to all five questions	0.62	0.70
=1 if YES to questions 1,2, and 5	0.84	0.88
Observations	1,737	1,261

Note: These five questions were asked in ACASI to everyone included in ACASI measurement experiment.

Table 3: Effect of Self-interviewing (SI) on Placebo Questions

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Malawi				
SI × Screen Pass (β)	-0.01	0.08***	0.09***	0.10***
	(0.01)	(0.03)	(0.03)	(0.03)
SI × Non-Pass (γ)	-0.05	0.11*	0.03	0.19***
	(0.04)	(0.06)	(0.07)	(0.06)
Screen Pass	0.02	-0.02	-0.01	0.03
	(0.03)	(0.05)	(0.06)	(0.05)
FTFI × Non-Pass mean	0.92	0.45	0.54	0.23
p-value ($\beta = \gamma$)	0.341	0.687	0.409	0.199
Observations	1,718	1,713	1,345	1,228
Effect of SI for the average respondent				
Pooled SI effects	-0.02	0.09	0.08	0.11
p-value	0.160	0.000	0.003	0.000
Panel B. Liberia				
SI × Screen Pass (β)	0.02	0.07***	-0.03**	-0.01
	(0.02)	(0.03)	(0.01)	(0.03)
SI × Non-Pass (γ)	-0.01	0.17**	-0.08*	-0.02
	(0.07)	(0.07)	(0.04)	(0.08)
Screen Pass	0.09	0.10*	0.02	0.03
	(0.06)	(0.06)	(0.03)	(0.06)
FTFI × Non-Pass mean	0.72	0.57	0.96	0.65
p-value ($\beta = \gamma$)	0.719	0.219	0.273	0.884
Observations	1,259	1,260	1,226	1,101
Effect of SI for the average respondent				
Pooled SI effects	0.01	0.09	-0.04	-0.01
p-value	0.496	0.001	0.002	0.787

Note: "Screen Pass" is defined by selecting "yes" to questions 1, 2, and 5 in Table 2. By this alternative definition, 84% in Malawi and 88% in Liberia are in the "Screen Pass" group. Regressions are at the respondent-question level. Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table 4: Effect of Self-interviewing (SI) on IPV (Individual Questions)

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *individual* question in the following category:				All questions pooled
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI	0.09***	0.05***	0.01*	0.03***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
FTFI mean	0.11	0.07	0.03	0.04	0.07
Number of individuals	1,715	1,711	1,712	1,709	1,716
Observations	11,887	6,802	10,181	5,095	33,965
Panel B. Liberia					
SI	0.05***	0.01	0.01	0.03***	0.03**
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
FTFI mean	0.20	0.19	0.09	0.04	0.14
Number of individuals	1,259	1,259	1,259	1,259	1,259
Observations	8,752	5,006	7,508	3,758	25,024

Note: Regressions are at the respondent-*question* level (violence is not aggregated into indexes). See Table 5 for results in which IPV questions are aggregated into indices. Regressions include question-level fixed effects and individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

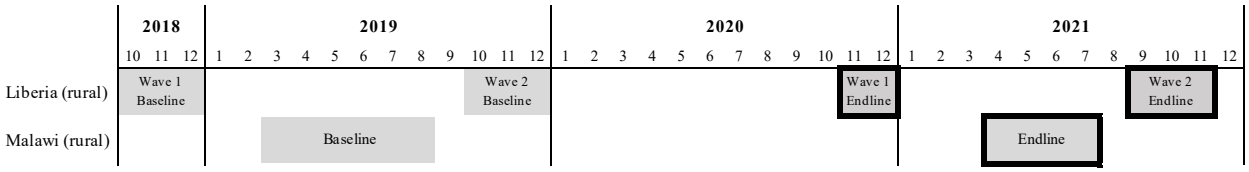
Table 5: Effect of Self-interviewing (SI) on IPV Indices

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *at least one* question in the following category:				Any IPV
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI	0.18***	0.10***	0.05***	0.06***	0.13***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
FTFI mean	0.38	0.16	0.08	0.07	0.20
Observations	1,737	1,737	1,737	1,737	1,737
Panel B. Liberia					
SI	0.07***	0.04	0.01	0.08***	0.04
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
FTFI mean	0.57	0.34	0.23	0.07	0.38
Observations	1,261	1,261	1,261	1,261	1,261

Note: IPV measures are indexed by category; index is set equal to 1 if the respondent answered "yes" to any question in the category. Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Appendix A: Additional Figures and Tables

Figure A1: Timeline of Survey Activities



Note: Bold rectangles refer to the survey rounds where ACASI vs. FTFI randomization was implemented. Liberia Wave 1 sample is excluded from our results in this paper, as most ACASI protocols were developed, tested, and refined during Liberia's Wave 1 Endline.

Figure A2: Self-interviewing Module

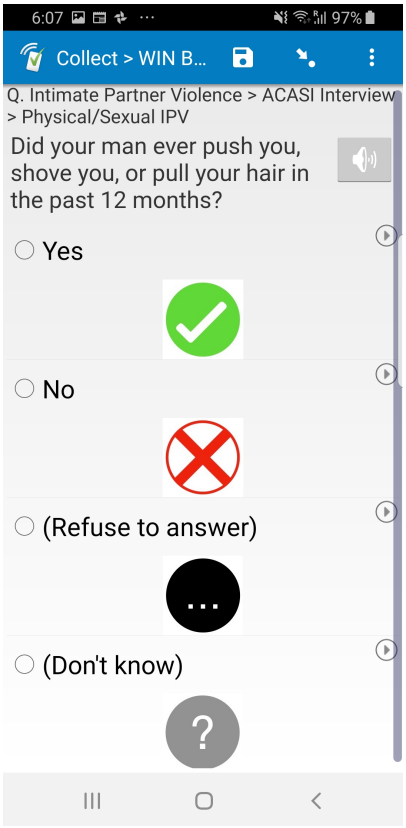
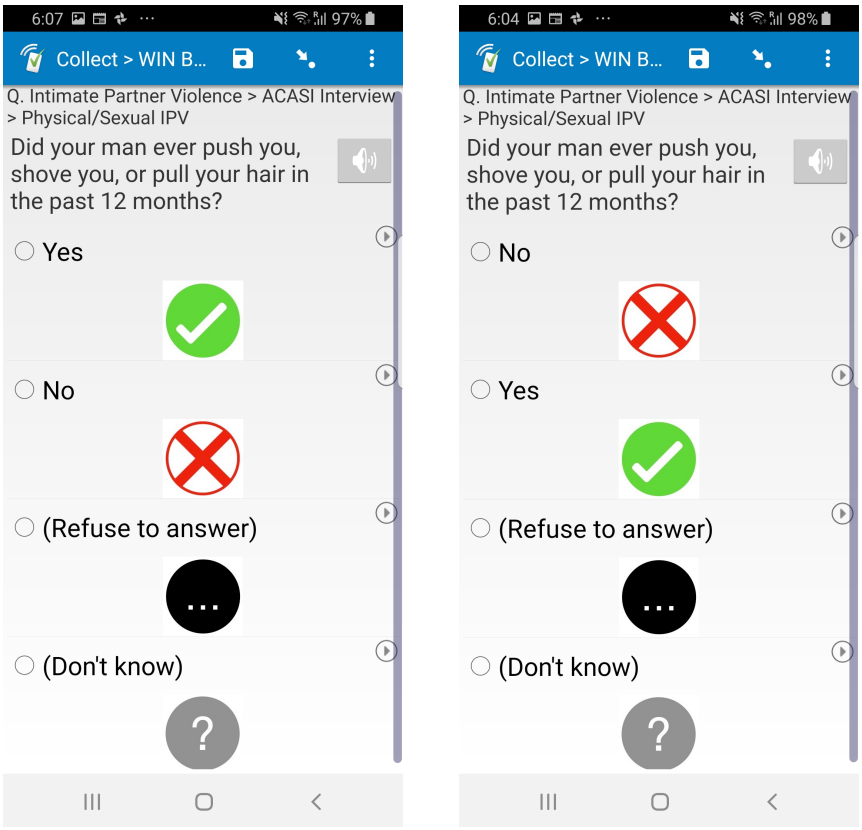


Figure A3: Appearance of Module with "yes" or "no" option appearing first



Notes: Women would see either the display on the left or right.

Table A1: Balance Check in Sample Exclusion

	(1) Refused to all IPV questions	(2) Dropped due to vision/hearing impairment
Experimentally assigned to ACASI	0.001	-0.001
	(0.002)	(0.009)
FTFI mean	0.004	0.068
Observations	3,220	3,220

Note: Observations pooled between Liberia and Malawi samples. Regressions include country sample fixed effects. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A2: Correlates of "Passing" SI Screening Questions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	=1 if passed SI screening[a]							
	Malawi				Liberia			
Years of education	0.017***			0.020***	0.001			0.005
	(0.002)			(0.004)	(0.003)			(0.004)
=1 if able to write/read		0.091***		-0.004		-0.013		-0.035
		(0.020)		(0.031)		(0.021)		(0.031)
=1 if has access to mobile phone			0.014	-0.006			0.004	-0.007
			(0.018)	(0.018)			(0.019)	(0.020)
R-square	0.024	0.014	0.000	0.033	0.000	0.000	0.000	0.023
Overall mean of outcome	0.84	0.84	0.84	0.84	0.88	0.88	0.88	0.88
Observations	1,737	1,737	1,737	1,737	1,261	1,261	1,261	1,261

Note: Columns 1-3 and 5-7 present bivariate regressions. Columns 4 and 8 include all variables in Table 1, but other coefficients are not reported for space. Robust standard error in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.
[a] Passing threshold is choosing "yes" for all the five questions in Table 2.

Table A3: Effect of SI on Choosing "Don't know" or "Refuse to answer" in Placebo Questions

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Malawi				
SI	0.013***	0.013**	-0.080***	-0.000
	(0.005)	(0.005)	(0.020)	(0.022)
FTFI mean	0.005	0.007	0.266	0.294
Observations	1,737	1,737	1,737	1,737
Panel B. Liberia				
SI	0.003	0.002	0.017*	-0.011
	(0.002)	(0.002)	(0.009)	(0.019)
FTFI mean	0.000	0.000	0.020	0.132
Observations	1,261	1,261	1,261	1,261

Note: Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A4: Effect of Self-interviewing (SI) on Placebo Questions, Alternative Definition of Passing

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Malawi				
SI × Screen Pass (β)	-0.01	0.08***	0.09***	0.10***
	(0.01)	(0.03)	(0.03)	(0.03)
SI × Non-Pass (γ)	-0.06*	0.10*	0.05	0.18***
	(0.03)	(0.06)	(0.06)	(0.06)
Screen Pass	0.00	-0.02	-0.03	0.04
	(0.02)	(0.04)	(0.05)	(0.04)
FTFI × Non-Pass mean	0.94	0.45	0.55	0.22
p-value ($\beta = \gamma$)	0.164	0.778	0.625	0.219
Observations	1,718	1,713	1,345	1,228
Effect of SI for the average respondent				
Pooled SI effects	-0.02	0.09	0.08	0.11
p-value	0.152	0.000	0.003	0.000
Panel B. Liberia				
SI × Screen Pass (β)	0.01	0.07**	-0.03**	-0.00
	(0.02)	(0.03)	(0.01)	(0.03)
SI × Non-Pass (γ)	0.02	0.13**	-0.06**	-0.02
	(0.05)	(0.05)	(0.03)	(0.06)
Screen Pass	0.06	0.01	0.01	-0.00
	(0.04)	(0.05)	(0.02)	(0.05)
FTFI × Non-Pass mean	0.76	0.65	0.97	0.70
p-value ($\beta = \gamma$)	0.896	0.306	0.298	0.796
Observations	1,259	1,260	1,226	1,101
Effect of SI for the average respondent				
Pooled SI effects	0.01	0.08	-0.04	-0.01
p-value	0.533	0.001	0.002	0.769

Note: Alternatively "Screen Pass" is defined by selecting "yes" to questions 1, 2, and 5 in Table 2 and also choosing "no" to the question "Have you traveled outside the country in the past week?" By this alternative definition, 81% in Malawi and 78% in Liberia are in the "Screen Pass" group. Regressions are at the respondent-question level. Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A5: Heterogeneity in Effects of ACASI on Placebo Questions (Malawi)

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Primary education completion				
SI × Primary Educ (β)	-0.02	0.10***	0.06	0.07*
	(0.01)	(0.03)	(0.04)	(0.04)
SI × No Primary Educ (γ)	-0.01	0.07**	0.10***	0.16***
	(0.02)	(0.03)	(0.04)	(0.03)
Primary Educ	0.03	0.03	0.04	0.10*
	(0.02)	(0.05)	(0.06)	(0.05)
FTFI × No Primary Educ mean	0.93	0.41	0.52	0.22
p-value ($\beta = \gamma$)	0.749	0.659	0.404	0.096
Observations	1,718	1,713	1,345	1,228
Panel B. Access to mobile phone				
SI × Mobile (β)	-0.01	0.13***	0.09**	0.09**
	(0.02)	(0.03)	(0.04)	(0.04)
SI × No Mobile (γ)	-0.02	0.05	0.07*	0.14***
	(0.02)	(0.03)	(0.04)	(0.03)
Mobile	-0.02	-0.02	-0.01	0.00
	(0.02)	(0.03)	(0.04)	(0.04)
FTFI × No Mobile mean	0.95	0.46	0.53	0.26
p-value ($\beta = \gamma$)	0.549	0.096	0.763	0.314
Observations	1,718	1,713	1,345	1,228
Panel C. Able to read/write				
SI × English (β)	-0.02*	0.10***	0.06*	0.10***
	(0.01)	(0.03)	(0.03)	(0.03)
SI × No English (γ)	-0.01	0.06	0.12***	0.15***
	(0.02)	(0.04)	(0.05)	(0.04)
English	0.03	-0.04	0.01	-0.00
	(0.02)	(0.05)	(0.05)	(0.05)
FTFI × No English mean	0.92	0.40	0.51	0.23
p-value ($\beta = \gamma$)	0.499	0.385	0.239	0.334
Observations	1,718	1,713	1,345	1,228
Panel D. Age				
SI × Below-median Age (β)	-0.03*	0.07**	0.04	0.10***
	(0.02)	(0.03)	(0.04)	(0.04)
SI × Above-median Age (γ)	-0.00	0.11***	0.12***	0.13***
	(0.02)	(0.03)	(0.04)	(0.04)
Below-median Age	-0.00	0.05	0.05	0.09*
	(0.02)	(0.05)	(0.05)	(0.05)
FTFI × Above-median Age mean	0.94	0.41	0.52	0.27
p-value ($\beta = \gamma$)	0.188	0.397	0.120	0.511
Observations	1,718	1,713	1,345	1,228

Note: Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A6: Heterogeneity in Effects of ACASI on Placebo Questions (Liberia)

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Primary education completion				
SI × Primary Educ (β)	0.07	0.00	-0.01	-0.06
	(0.05)	(0.05)	(0.03)	(0.06)
SI × No Primary Educ (γ)	-0.00	0.11***	-0.04***	0.01
	(0.02)	(0.03)	(0.01)	(0.03)
Primary Educ	-0.11*	0.01	-0.03	-0.06
	(0.06)	(0.07)	(0.03)	(0.07)
FTFI × No Primary Educ mean	0.82	0.66	0.98	0.68
p-value ($\beta = \gamma$)	0.159	0.097	0.194	0.337
Observations	1,259	1,260	1,226	1,101
Panel B. Access to mobile phone				
SI × Mobile (β)	0.08**	0.05	-0.03*	0.02
	(0.03)	(0.04)	(0.02)	(0.04)
SI × No Mobile (γ)	-0.03	0.11***	-0.04***	-0.03
	(0.03)	(0.03)	(0.02)	(0.04)
Mobile	-0.07**	0.09**	0.00	-0.00
	(0.03)	(0.04)	(0.01)	(0.04)
FTFI × No Mobile mean	0.83	0.63	0.98	0.69
p-value ($\beta = \gamma$)	0.017	0.293	0.677	0.313
Observations	1,259	1,260	1,226	1,101
Panel C. Able to read/write				
SI × English (β)	0.02	0.04	-0.02	0.01
	(0.04)	(0.05)	(0.02)	(0.05)
SI × No English (γ)	0.01	0.11***	-0.04***	-0.02
	(0.03)	(0.03)	(0.01)	(0.03)
English	0.00	0.03	-0.01	0.01
	(0.04)	(0.05)	(0.02)	(0.05)
FTFI × No English mean	0.81	0.66	0.98	0.69
p-value ($\beta = \gamma$)	0.861	0.203	0.512	0.597
Observations	1,259	1,260	1,226	1,101
Panel D. Age				
SI × Below-median Age (β)	-0.01	0.09**	-0.02	0.01
	(0.03)	(0.04)	(0.02)	(0.04)
SI × Above-median Age (γ)	0.03	0.08**	-0.05***	-0.02
	(0.03)	(0.04)	(0.02)	(0.04)
Below-median Age	0.04	-0.03	-0.01	0.04
	(0.04)	(0.05)	(0.02)	(0.05)
FTFI × Above-median Age mean	0.82	0.68	0.98	0.70
p-value ($\beta = \gamma$)	0.399	0.816	0.277	0.607
Observations	1,259	1,260	1,226	1,101

Note: Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A7: Attention Span and Placebo Effects (Malawi)

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Survey time of the day				
SI × Afternoon (β)	-0.02	0.11***	0.08***	0.12***
	(0.01)	(0.03)	(0.03)	(0.03)
SI × Morning (γ)	-0.02	-0.02	0.07	0.08
	(0.03)	(0.05)	(0.06)	(0.06)
Afternoon	-0.01	-0.12***	-0.02	-0.04
	(0.02)	(0.04)	(0.05)	(0.05)
FTFI × Morning mean	0.96	0.55	0.54	0.30
p-value ($\beta = \gamma$)	0.808	0.031	0.930	0.478
Observations	1,715	1,710	1,343	1,228
Panel B. Food security index				
SI × Higher food security (β)	-0.05***	0.10***	-0.05	0.08**
	(0.02)	(0.03)	(0.04)	(0.04)
SI × Lower food security (γ)	0.01	0.07**	0.20***	0.15***
	(0.01)	(0.03)	(0.04)	(0.03)
Higher food security	0.04*	-0.03	0.15***	0.05
	(0.02)	(0.05)	(0.05)	(0.05)
FTFI × Lower food security mean	0.94	0.42	0.40	0.17
p-value ($\beta = \gamma$)	0.013	0.559	0.000	0.192
Observations	1,718	1,713	1,345	1,228

Note: Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A8: Attention Span and Placebo Effects (Liberia)

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Survey time of the day				
SI × Afternoon (β)	0.06*	0.03	-0.01	0.02
	(0.04)	(0.04)	(0.02)	(0.05)
SI × Morning (γ)	-0.02	0.10***	-0.06***	-0.04
	(0.03)	(0.04)	(0.02)	(0.04)
Afternoon	-0.05	0.03	-0.02*	-0.06
	(0.03)	(0.04)	(0.01)	(0.04)
FTFI × Morning mean	0.82	0.65	0.99	0.72
p-value ($\beta = \gamma$)	0.069	0.243	0.046	0.334
Observations	1,126	1,127	1,096	988
Panel B. Food security index				
SI × Higher food security (β)	0.04	0.11***	-0.02*	0.04
	(0.03)	(0.03)	(0.01)	(0.04)
SI × Lower food security (γ)	-0.02	0.05	-0.05***	-0.06
	(0.03)	(0.04)	(0.02)	(0.04)
Higher food security	-0.03	0.01	-0.02	-0.02
	(0.04)	(0.05)	(0.02)	(0.05)
FTFI × Lower food security mean	0.79	0.66	0.97	0.65
p-value ($\beta = \gamma$)	0.187	0.224	0.261	0.067
Observations	1,259	1,260	1,226	1,101

Note: Regressions include individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A9: Does the effect of ACASI differ between those who pass screening and those who don't? (Individual IPV Questions)

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *individual* question in the following category:				All questions pooled
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI × Screen Pass (β)	0.08***	0.04***	0.01	0.03***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SI × Non-Pass (γ)	0.13***	0.12***	0.05**	0.06**	0.09***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Screen Pass	-0.02	0.01	0.01	0.01	0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
FTFI × Non-Pass mean	0.13	0.06	0.03	0.03	0.07
p-value ($\beta = \gamma$)	0.140	0.014	0.052	0.169	0.021
Number of individuals	1,715	1,711	1,712	1,709	1,716
Observations	11,887	6,802	10,181	5,095	33,965
Effect of SI for the average respondent					
Pooled SI effects	0.09	0.05	0.01	0.03	0.05
p-value	0.000	0.000	0.077	0.001	0.000
Panel B. Liberia					
SI × Screen Pass (β)	0.04***	0.01	0.00	0.02**	0.02*
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
SI × Non-Pass (γ)	0.12***	0.03	0.04	0.09**	0.07**
	(0.04)	(0.05)	(0.04)	(0.04)	(0.03)
Screen Pass	0.04	0.00	-0.00	-0.00	0.02
	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)
FTFI × Non-Pass mean	0.18	0.20	0.10	0.05	0.14
p-value ($\beta = \gamma$)	0.084	0.749	0.379	0.091	0.168
Number of individuals	1,259	1,259	1,259	1,259	1,259
Observations	8,752	5,006	7,508	3,758	25,024
Effect of SI for the average respondent					
Pooled SI effects	0.05	0.01	0.01	0.03	0.03
p-value	0.000	0.485	0.451	0.006	0.013

Note: Regressions are at the respondent-*question* level (violence is not aggregated into indexes). Regressions include question-level fixed effects and individual controls (including all variables in Table 1). "Screen Pass" is defined by selecting "yes" to all questions in Table 2. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A10: Effect of ACASI on Choosing "Don't know" or "Refuse to answer" in IPV Questions

	(1)	(2)	(3)	(4)	(5)
	=1 if don't know or refusal to *individual* question in following category:				All questions pooled
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI	0.011**	0.007*	0.011**	0.006**	0.012**
	(0.006)	(0.003)	(0.005)	(0.003)	(0.006)
FTFI mean	0.017	0.009	0.015	0.007	0.017
Number of individuals	1,737	1,737	1,737	1,737	1,737
Observations	12,159	12,159	12,159	12,159	34,740
Panel B. Liberia					
SI	0.007**	0.002	-0.001	0.003**	0.004
	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)
FTFI mean	0.005	0.004	0.007	0.001	0.006
Number of individuals	1,261	1,261	1,261	1,261	1,261
Observations	8,827	8,827	8,827	8,827	25,220

Note: Regressions are at the respondent-*question* level (violence is not aggregated into indexes). Regressions include question-level fixed effects and individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Appendix B: Self-reported technical and comprehension difficulties

Table B1: Debriefing Survey on Technical Issues with ACASI Module

	(1)	(2)
	Mean (=1 if yes)	
	Malawi	Liberia
Was the audio loud enough to hear?	0.99	0.98
Was the audio speaking speed okay?	0.99	0.98
Observations	866	616

Note: Questions were asked only to those in the ACASI treatment group (i.e., the FTFI group did not get these questions).

Table B2: Relationship between Reporting Technical Difficulties and Passing Screening

	(1)	(2)	(3)	(4)
	=1 if passed SI screening[a]			
	Malawi		Liberia	
=1 if said:				
audio loud enough to hear	-0.091		0.036	
	(0.081)		(0.109)	
audio speaking speed okay		0.026		0.036
		(0.180)		(0.109)
R-square	0.001	0.000	0.000	0.000
Outcome mean when said no	0.92	0.80	0.83	0.83
Observations	866	867	616	615

Note: Robust standard errors are in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.
[a] Passing threshold is choosing "yes" for questions 1,2, and 5 in Table 2.

Table B3: Debriefing Survey on Comprehension Issues with ACASI Module

	(1)	(2)
	Mean (=1 if yes)	
	Malawi	Liberia
Was it easy for you to remember the meaning of pictures?	0.90	0.90
Was it easy for you to choose answers on the screen?	0.91	0.88
Was it easy for you to move between questions on the screen?	0.92	0.88
Observations	866	616

Note: Questions were asked only to those in the ACASI treatment group (i.e., the FTFI group did not get these questions).

Table B4: Relationship between Reporting Comprehension and Passing Screening

	(1)	(2)	(3)	(4)	(5)	(6)
	=1 if passed SI screening[a]					
	Malawi			Liberia		
=1 if said:						
easy to remember the meaning of pictures	0.117**			0.101*		
	(0.049)			(0.054)		
easy to choose answers on screen		0.054			0.068	
		(0.048)			(0.048)	
easy to move between questions on screen			0.057			0.075
			(0.051)			(0.048)
R-square	0.009	0.002	0.002	0.008	0.004	0.005
Outcome mean when said no	0.72	0.78	0.77	0.78	0.81	0.80
Observations	866	865	866	616	616	616

Note: Robust standard errors are in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.
[a] Passing threshold is choosing "yes" for questions 1, 2, and 5 in Table 2.

Appendix C: Subtreatments

Table C1: Effect of Ordering of Yes and No Options in ACASI on Placebo Questions

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Malawi				
YES First	-0.02	0.05	0.06*	0.07*
	(0.02)	(0.03)	(0.04)	(0.04)
NO First mean	0.94	0.54	0.58	0.34
Observations	854	851	708	615
Panel B. Liberia				
YES First	0.00	0.03	0.01	-0.01
	(0.03)	(0.04)	(0.02)	(0.04)
NO First mean	0.82	0.74	0.94	0.70
Observations	615	616	595	542

Note: Includes only those who are in the ACASI group (FTFI group excluded). Regressions include country sample fixed effects and individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table C2: Effect of Ordering of Yes and No Options in ACASI on IPV Questions

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *individual* question in the following category:				All questions pooled
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
YES First	-0.01	0.00	0.00	-0.00	-0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
NO First mean	0.21	0.13	0.05	0.07	0.12
Number of individuals	858	854	855	852	859
Observations	5,915	3,385	5,062	2,531	16,893
Panel B. Liberia					
YES First	-0.01	-0.02	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
NO First mean	0.25	0.19	0.10	0.08	0.17
Number of individuals	617	617	617	617	617
Observations	4,268	2,446	3,676	1,832	12,222

Note: Includes only those who are in the ACASI group (FTFI group excluded). Observations at respondent-question level. Regressions include question-level fixed effects and individual controls (including all variables in Table 1). Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table C3: Effect of Placebo Module Position on SI Effects for Placebo Questions

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Malawi				
SI × Placebos First (β)	0.01	0.08**	0.11***	0.11***
	(0.02)	(0.03)	(0.04)	(0.04)
SI × IPV First (γ)	-0.04**	0.08***	0.05	0.12***
	(0.02)	(0.03)	(0.04)	(0.03)
Placebos First	0.00	0.00	0.01	0.05
	(0.02)	(0.03)	(0.04)	(0.03)
FTFI × IPV First mean	0.94	0.46	0.54	0.26
p-value ($\beta = \gamma$)	0.036	0.992	0.326	0.895
Observations	1,718	1,713	1,345	1,228
Panel B. Liberia				
SI × Placebos First (β)	0.03	0.10***	-0.03*	0.01
	(0.03)	(0.04)	(0.02)	(0.04)
SI × IPV First (γ)	-0.01	0.07**	-0.04**	-0.03
	(0.03)	(0.04)	(0.02)	(0.04)
Placebos First	-0.01	-0.04	-0.01	-0.04
	(0.03)	(0.04)	(0.01)	(0.04)
FTFI × IPV First mean	0.81	0.69	0.98	0.70
p-value ($\beta = \gamma$)	0.356	0.633	0.610	0.446
Observations	1,259	1,260	1,226	1,101

Note: Regressions include individual controls (including all variables in Table 1) and the order between IPV module and PHQ-9 module. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table C4: Effect of Placebo Module Position on SI Effects for IPV Questions

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *individual* question in the following category:				All questions pooled
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI × Placebos First (β)	0.10***	0.05***	0.02**	0.05***	0.06***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
SI × IPV First (γ)	0.08***	0.05***	0.00	0.01	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Placebos First	0.01	0.02	-0.01	-0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
FTFI × IPV First mean	0.11	0.06	0.04	0.04	0.07
p-value ($\beta = \gamma$)	0.275	0.916	0.276	0.030	0.217
Observations	11,887	6,802	10,181	5,095	33,965
Panel B. Liberia					
SI × Placebos First (β)	0.05***	-0.00	0.00	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
SI × IPV First (γ)	0.05**	0.02	0.01	0.04**	0.03**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Placebos First	0.00	0.00	0.01	0.00	0.01
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
FTFI × IPV First mean	0.20	0.19	0.09	0.04	0.14
p-value ($\beta = \gamma$)	0.965	0.415	0.587	0.415	0.614
Observations	8,752	5,006	7,508	3,758	25,024

Note: Observations at respondent-question level. Regressions include question-level fixed effects, individual controls (including all variables in Table 1), and the order between IPV module and PHQ-9 module. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Appendix D: Main results, without controls

Table D1: Effect of Self-interviewing (SI) on Placebo Questions, no individual controls

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Malawi				
SI × Screen Pass (β)	-0.01	0.07***	0.07**	0.06**
	(0.01)	(0.03)	(0.03)	(0.03)
SI × Non-Pass (γ)	-0.05	0.13**	0.03	0.19***
	(0.04)	(0.06)	(0.07)	(0.07)
Screen Pass	0.02	0.02	0.01	0.07
	(0.03)	(0.05)	(0.06)	(0.05)
FTFI × Non-Pass mean	0.92	0.45	0.54	0.23
p-value ($\beta = \gamma$)	0.285	0.365	0.622	0.069
Observations	1,718	1,713	1,345	1,228
Effect of SI for the average respondent				
Pooled SI effects	-0.02	0.08	0.06	0.09
p-value	0.180	0.001	0.019	0.001
Panel B. Liberia				
SI × Screen Pass (β)	0.02	0.08***	-0.03**	0.00
	(0.02)	(0.03)	(0.01)	(0.03)
SI × Non-Pass (γ)	0.00	0.17**	-0.08*	-0.02
	(0.07)	(0.08)	(0.04)	(0.08)
Screen Pass	0.10*	0.11*	0.02	0.05
	(0.05)	(0.06)	(0.02)	(0.06)
FTFI × Non-Pass mean	0.72	0.57	0.96	0.65
p-value ($\beta = \gamma$)	0.862	0.248	0.290	0.846
Observations	1,259	1,260	1,226	1,101
Effect of SI for the average respondent				
Pooled SI effects	0.02	0.09	-0.03	-0.00
p-value	0.451	0.001	0.002	0.967

Note: Regressions are at the respondent-question level. "Screen Pass" is defined by selecting "yes" to the first five questions in Table 2. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table D2: Effect of Self-interviewing (SI) on IPV Reporting in Individual Questions, no individual controls

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *individual* question in the following category:				All questions pooled
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI	0.09***	0.06***	0.01*	0.03***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
FTFI mean	0.11	0.07	0.03	0.04	0.07
Number of individuals	1,715	1,711	1,712	1,709	1,716
Observations	11,887	6,802	10,181	5,095	33,965
Panel B. Liberia					
SI	0.05***	0.00	0.00	0.03***	0.02**
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
FTFI mean	0.20	0.19	0.09	0.04	0.14
Number of individuals	1,259	1,259	1,259	1,259	1,259
Observations	8,752	5,006	7,508	3,758	25,024

Note: Regressions are at the respondent-*question* level (violence is not aggregated into indexes). See Table D3 for results in which IPV questions are aggregated into indices. Regressions include question-level fixed effects. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table D3: Effect of Self-interviewing (SI) on IPV Indices, no individual controls

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *at least one* question in the following category:				Any IPV
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI	0.18***	0.11***	0.05***	0.06***	0.14***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
FTFI mean	0.38	0.16	0.08	0.07	0.20
Observations	1,737	1,737	1,737	1,737	1,737
Panel B. Liberia					
SI	0.07**	0.03	0.01	0.07***	0.03
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
FTFI mean	0.57	0.34	0.23	0.07	0.38
Observations	1,261	1,261	1,261	1,261	1,261

Note: IPV measures are indexed by category; index is set equal to 1 if the respondent answered "yes" to any question in the category. "Screen Pass" is defined by selecting "yes" to all questions in Table 2. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Appendix E: Main results, unconditional on "don't know" or "refuse to answer"

Table E1: Effect of Self-interviewing (SI) on Placebo Questions, unconditional sample

	(1) Farm work (past year)	(2) Market visit (past week)	(3) Maize/Rice (next week)	(4) Meat (next week)
Panel A. Malawi				
SI × Screen Pass (β)	-0.02	0.08***	0.12***	0.06***
	(0.01)	(0.03)	(0.03)	(0.02)
SI × Non-Pass (γ)	-0.08*	0.08	0.00	0.12**
	(0.04)	(0.06)	(0.06)	(0.05)
Screen Pass	0.05	-0.01	-0.03	0.02
	(0.03)	(0.05)	(0.05)	(0.04)
FTFI × Non-Pass mean	0.89	0.44	0.41	0.17
p-value ($\beta = \gamma$)	0.142	0.995	0.066	0.264
Observations	1,737	1,737	1,737	1,737
Effect of SI for the average respondent				
Pooled SI effects	-0.03	0.08	0.10	0.07
p-value	0.030	0.001	0.000	0.001
Panel B. Liberia				
SI × Screen Pass (β)	0.02	0.07***	-0.04***	0.00
	(0.02)	(0.03)	(0.01)	(0.03)
SI × Non-Pass (γ)	-0.02	0.17**	-0.09*	0.00
	(0.07)	(0.07)	(0.05)	(0.08)
Screen Pass	0.09	0.10*	0.04	0.02
	(0.06)	(0.06)	(0.03)	(0.06)
FTFI × Non-Pass mean	0.72	0.57	0.92	0.56
p-value ($\beta = \gamma$)	0.645	0.212	0.361	0.968
Observations	1,261	1,261	1,261	1,261
Effect of SI for the average respondent				
Pooled SI effects	0.01	0.08	-0.05	0.00
p-value	0.576	0.001	0.001	0.957

Note: Observations for which people responded "don't know" or refused to answer are imputed with zeroes, instead of being dropped. Regressions are at the respondent-question level. "Screen Pass" is defined by selecting "yes" to the first five questions in Table 2. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table E2: Effect of Self-interviewing (SI) on IPV Reporting in Individual Questions, unconditional sample

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *individual* question in the following category:				All questions pooled
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI	0.09***	0.05***	0.01*	0.03***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
FTFI mean	0.11	0.07	0.03	0.04	0.07
Number of individuals	1,737	1,737	1,737	1,737	1,737
Observations	12,159	6,948	10,422	5,211	34,740
Panel B. Liberia					
SI	0.05***	0.01	0.01	0.03***	0.03**
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
FTFI mean	0.20	0.19	0.09	0.04	0.14
Number of individuals	1,261	1,261	1,261	1,261	1,261
Observations	8,827	5,044	7,566	3,783	25,220

Note: Observations for which people responded "don't know" or refused to answer are imputed with zeroes, instead of being dropped. Regressions are at the respondent-*question* level (violence is not aggregated into indexes). See Table D3 for results in which IPV questions are aggregated into indices. Regressions include question-level fixed effects. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table E3: Effect of Self-interviewing (SI) on IPV Indices, unconditional sample

	(1)	(2)	(3)	(4)	(5)
	=1 if responded yes to *at least one* question in the following category:				Any IPV
	Controlling Behavior	Emotional IPV	Physical IPV	Sexual IPV	
Panel A. Malawi					
SI	0.18***	0.10***	0.05***	0.06***	0.13***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
FTFI mean	0.38	0.16	0.08	0.07	0.20
Observations	1,737	1,737	1,737	1,737	1,737
Panel B. Liberia					
SI	0.07***	0.04	0.01	0.08***	0.04
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
FTFI mean	0.57	0.34	0.23	0.07	0.38
Observations	1,261	1,261	1,261	1,261	1,261

Note: Observations for which people responded "don't know" or refused to answer are imputed with zeroes, instead of being dropped. IPV measures are indexed by category; index is set equal to 1 if the respondent answered "yes" to any question in the category. "Screen Pass" is defined by selecting "yes" to all questions in Table 2. Robust standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Appendix F: Survey instrument

Non-sensitive placebo questions[23]

1. Did you do any farm work in the past year?
2. Did you go to the market in the past week?
3. Will you, or anyone in your household, eat any [rice/maize] next week, one time or more?
4. Will you, or anyone in your household, eat any type of meat next week, one time or more?

Controlling behavior

1. Did your man ever try to keep you from seeing your friends in the past 12 months?
2. Did your man ever try to stop you from meeting or speaking to your family of birth in the past 12 months?
3. Did your man ever need to know where you are all the time in the past 12 months?
4. Did your man ever stop talking to you or treat you with no interest in the past 12 months?
5. Did your man ever get angry if you speak with another man in the past 12 months?
6. Did your man often think that you are unfaithful in the past 12 months?
7. In the past 12 months, did your man ever expect you to ask for his approval before you go to a health clinic or hospital?

Emotional IPV[24]

1. Did your man ever insult you or make you feel bad about yourself in the past 12 months?
2. Did your man ever make you feel small in front of other people in the past 12 months?
3. Did your man ever mean to scare you (for example, by the way he looked at you, by yelling and bursting things) in the past 12 months?
4. Did your man ever threaten to hurt you or someone you care about in the past 12 months?

Physical IPV[24]

1. Did your man ever slap you or throw something at you that could hurt you in the past 12 months?
2. Did your man ever push you, shove you, or pull your hair in the past 12 months?
3. Did your man ever hit you with his hand or with something else that could hurt you in the past 12 months?

[23] Some questions asked in the ACASI Screening module were re-asked later either by SI or by FTFI. For placebo effects analysis, we exclude those questions and include only the four questions listed here, which were not previously asked in the screening module.

[24] For each IPV question, if the answer is "yes", a follow-up question about frequency appears, asking whether it happened (i) one or two times, (ii) three to five times, or (iii) more than five times.

4. Did your man ever kick you, drag you or beat you up in the past 12 months?
5. Did your man ever mean to choke or burn you in the past 12 months?
6. Did your man ever threaten to use or actually use a gun, knife or other weapon against you in the past 12 months?

Sexual IPV[24]

1. Did your man ever physically force you to do man and woman business when you did not want to in the past 12 months?
2. Did you ever do man and woman business when you did not want to because you were afraid of what your man might do in the past 12 months?
3. In the past 12 months, while doing man and woman business, did your man ever force you to do something that made you feel small or bad about yourself?