IFPRI Discussion Paper 02094

December 2021

Using a List Experiment to Measure Intimate Partner Violence

Cautionary Evidence from Ethiopia

Daniel O. Gilligan

Melissa Hidrobo

Jessica Leight

Heleene Tambet

Poverty, Health, and Nutrition Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), a CGIAR Research Center established in 1975, provides research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition. IFPRI's strategic research aims to foster a climate-resilient and sustainable food supply; promote healthy diets and nutrition for all; build inclusive and efficient markets, trade systems, and food industries; transform agricultural and rural economies; and strengthen institutions and governance. Gender is integrated in all the Institute's work. Partnerships, communications, capacity strengthening, and data and knowledge management are essential components to translate IFPRI's research from action to impact. The Institute's regional and country programs play a critical role in responding to demand for food policy research and in delivering holistic support for country-led development. IFPRI collaborates with partners around the world.

AUTHORS

Daniel O. Gilligan (d.gilligan@cgiar.org) is the Deputy Division Director in the Poverty, Health, and Nutrition Division of the International Food Policy Research Institute (IFPRI), Washington, DC..

Melissa Hidrobo (m.hidrobo@cgiar.org) is a Senior Research Fellow in IFPRI's Poverty, Health, and Nutrition Division, Washington, DC.

Jessica Leight (j.leight@cgiar.org) s a Research Fellow in IFPRI's Poverty, Health and Nutrition Division, Washington, DC.

Heleene Tambet (h.tambet@cgiar.org) is a Research Associate in IFPRI's Poverty, Health and Nutrition Division, Washington, DC.

Notices

[1] IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by IFPRI.

[2] The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

[3] Copyright remains with the authors. The authors are free to proceed, without further IFPRI permission, to publish this paper, or any revised version of it, in outlets such as journals, books, and other publications.

Abstract

While indirect methods are increasingly widely used to measure sensitive behaviors such as intimate partner violence in order to minimize social desirability biases in responses, in developing countries the use of more complex indirect questioning methods raises important questions around how individuals will react to the use of a more unusual and complex question structure. This paper presents evidence from a list experiment measuring multiple forms of intimate partner violence within an extremely poor sample of women in rural Ethiopia. We find that the list experiment does not generate estimates of intimate partner violence that are higher than direct response questions; rather, prevalence estimates using the list experiment are lower vis-à-vis prevalence estimates using the direct reports, and sometimes even negative. We interpret this finding as consistent with "fleeing" behavior by respondents who do not wish to be associated with statements associated with intimate partner violence.

Acknowledgments

We gratefully acknowledge funding for the impact evaluation by USAID under Cooperative Agreement No AID-FFP-A-16-00008. This work was undertaken as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by the International Food Policy Research Institute (IFPRI). Additional funding support for this study was also provided by PIM.

1. Introduction

Intimate partner violence (IPV) is a major global health challenge and is associated with a range of physical and psychological consequences for women and their families (Campbell, 2002; Devries et al., 2013; Ellsberg et al., 2008). Designing appropriate policies to reduce IPV requires high-quality data on its prevalence. However, measurement of sensitive and socially sanctioned behaviors such as IPV is challenging, typically entailing substantial investments in appropriate field protocols. Accordingly, researchers and policymakers hypothesize that underreporting of violence may remain significant, particularly in low-income settings where social norms can entail a normalization of intimate partner violence and a critical view toward women who report this behavior to outsiders.

In light of these challenges, field research in public health and economics has increasingly explored the use of list experiments and other indirect methods to measure IPV in large-scale surveys.[1] A list experiment is a strategy employed to measure the prevalence of a sensitive behavior (in this case, IPV) in a large sample, by randomizing respondents to receive a "control" list of non-sensitive "baseline" statements, or a "treatment" list including the baseline list and a statement capturing the sensitive behavior; respondents specify the number of statements that are true, and the difference in response means between the control and treatment groups allows the researcher to infer the prevalence of the sensitive behavior.[2] Importantly, this method does not require respondents to individually identify themselves as experiencing IPV. If social desirability bias meaningfully constrains respondents from directly reporting violence, the list experiment will then generate higher (and presumably more accurate estimates) of the sensitive behavior (Hinsley et al., 2019).

In this paper, we report on a large-scale survey conducted in rural Ethiopia that used standard direct elicitation methods as well as multiple list experiments to measure the prevalence of different dimensions of IPV (emotional, physical and sexual). Given that some forms of IPV may be more acceptable than others across settings, underreporting of IPV may depend on the type of violence. We find that the estimated rate of IPV using direct elicitation is low relative to other nationwide surveys; however, there is no evidence that the list experiment is effective in eliciting higher rates of IPV vis-à-vis the use of direct elicitation methods. Estimates of prevalence constructed using the list experiment are generally lower than the estimated prevalence using direct reporting methods (though the differences are not statistically significant), and in some cases are negative, especially for more sensitive forms of violence.

We interpret the low and even negative prevalence estimates constructed using the list experiment as evidence of respondents "fleeing" from concurrence with the IPV-linked item when presented with a list including this item. Perhaps due to lack of comprehension, respondents in fact seem more averse to associating themselves with IPV in a list experiment question than in a direct elicitation question. In addition, the "fleeing" response is more meaningful for more sensitive forms of IPV, for respondents who are characterized by high self-efficacy, and for control lists that do not include other sensitive items.

Relative to the existing literature, our results are in fact consistent with a growing body of evidence suggesting that list experiments may not be uniformly effective for measurement of IPV in developing country contexts, though the evidence pattern is mixed. An analysis of a household survey conducted in

[1] Hinsley et al. (2019) provides an overview of this literature.

[2] In some contexts, list experiments are described as item count methods.

Kerala, India used both direct reports and list experiments to collect data about lifetime experience of physical IPV, as well as physical harassment on public transportation. For IPV, indirect response methods generated a significantly higher estimated prevalence rate, while for physical harassment, the pattern was inverted (Joseph et al., 2017). Two other papers in Peru and in Ethiopia found similar prevalence rates of IPV when estimated with list experiments and direct questions (Agüero & Frisancho, 2021; Kotsadam & Villanger, 2019). In contrast, other recent papers found list experiments generated higher estimates of reported IPV vis-à-vis direct methods in Sri Lanka, Vietnam, Rwanda, Nigeria, Burkina Faso (Bulte & Lensink, 2019; Cullen, 2020; Lépine et al., 2020; Traunmuller et al., 2018). We contribute to this literature by focusing on three types of violence, emotional, physical, and sexual, and find no difference in prevalence across list or direct reports.

The evidence base is similarly mixed for measurement of other sensitive health and social behaviors.[3] The generally weaker performance of list experiments in these domains may reflect the salience of challenges that have been identified; in particular, the use of list experiments is valid only under certain assumptions that include no design, floor or ceiling effects, discussed in further detail below (Blair & Imai, 2012). List experiments also generate less precise estimates relative to direct elicitation methods, require large samples, and are generally not effective for the measurement of very rare behaviors given limited precision (Hinsley et al., 2019). Moreover, participants may not trust the list experiment and comprehension among low-literacy populations may be a challenge (Chuang et al., 2021; Hinsley et al., 2019). Given the anonymity in respondent behavior, the use of list experiment data to evaluate heterogeneity in the prevalence of the behavior with respect to other characteristics is also limited (Blair & Imai, 2012). In conjunction with the existing literature, our results suggest that more caution may be warranted in using list experiments to measure IPV and/or other sensitive behaviors in a developing country context, and we further show what respondent and list characteristics may lead to the breakdown of the list performance.

2. Survey design and sample

a. Survey design

We conducted a survey experiment between July and October 2019 in the regions of Oromia and Amhara, Ethiopia to test whether direct reporting or the list method was more effective at eliciting women's reports of IPV.[4] The base sample includes 4,062 rural households in 193 kebeles (subdistricts) in 13 woredas (districts) in these two regions, sampled as part of a larger evaluation.

Approximately 55 percent of women interviewed were randomized to receive the list questions, with the remaining 45 percent randomized to receive only direct report questions (no list questions). Randomization was conducted with stratification at the woreda level, and was implemented by the research team for the full sample list prior to the launch of the survey. Those who were randomly

[3] Other papers have shown that list experiments do not generate relatively higher estimates of prevalence (vis-à-vis direct questions) for HIV status (Haber et al., 2018), condom use and sexual activity (Chong et al., 2020), and abortion (Bell & Bishai, 2019; Chuang et al., 2021; Moseson et al., 2017), and there is only weak evidence the experiment generates differentially higher reporting of female genital mutilation (De Cao & Lutz, 2018) and multiple sexual partners (Chuang et al., 2021). List experiments were effective for eliciting higher reports of abortions in Liberia (Moseson et al., 2015), and lower estimates of prevalence of condom use among female sex workers in Senegal, correcting a presumed overreporting (Treibich & Lépine, 2019). This is in contrast to the political science literature, where list experiments are generally more effective than direct elicitation in generating accurate reports of sensitive behaviors (Blair et al., 2020).

[4] The survey experiment was part of a larger randomized control trial to estimate the impact of a livelihood and nutrition intervention on multiple outcomes including women's well-being.

assigned to receive the list questions were further randomized into one of four groups: group A received the list questions for emotional violence; group B received the list questions for physical violence; group C received the list questions for sexual violence; and Group D received the control list.

All respondents randomized to receive the list questions were exposed to three different control lists composed of 4 statements (full text of the list questions is provided in Appendix 1). The fifth statement depended on the randomized group. For example, those in treatment group A, received the three control lists, each with an additional statement related to emotional violence. The control lists were generally comprised of standard questions probing whether or not the respondent engaged in activities that were non-controversial, but also relatively infrequent: e.g., attending a wedding, traveling to visit a friend, or attending a funeral. However, two of the three control lists (list one and list two) also included one relatively sensitive question: list one posed a question about whether the respondent was familiar with a case of forced marriage, and list two posed a question about whether the respondent knew of a friend or relative's daughter forced to undergo female genital mutilation. List three, however, included no sensitive items.

The selection of control items was guided by recommendations in the literature: in particular, control items should be selected such that few respondents would respond affirmatively or negatively to all items in the list, thus avoiding one form of ceiling or floor effects in which many respondents respond affirmatively or negatively to all list items (Blair & Imai, 2012). In addition, the inclusion of some relatively sensitive items in the control list, even if posed with reference to members of the community or other contacts, serves to "camouflage" the researcher's intent in measuring IPV, a strategy recommended in the literature (Chuang et al., 2021).

All eligible women regardless of whether or not they received the list questions were additionally administered the direct IPV questions from the gold standard modified conflict tactic scale – as collected in the 2016 Ethiopia DHS. These questions included the three types of violence (emotional, physical and sexual) included in the list questions.

Our analysis also draws on data about other respondent characteristics. This includes the respondent's age and education; a score on the new general self-efficacy scale, a question that uses a ladder image to assess the relative degree of control that respondents report over their own lives[5]; and whether the respondent concurs that IPV can be justified in any one of a series of specified scenarios (if a woman goes out with telling her husband, neglects her children, argues with her husband, burns the food, or refuses to have sex with her husband).

b. Sample

The sample analyzed in this paper is the 55 percent of women who were randomized to receive both the list and the direct response IPV questions. The remaining 45% were, based on the original experimental design, randomized to receive only the direct response questions and were not targeted for the list experiments.[6] Of the 2,122 households randomized into the LIST group, 2,083 included a successfully completed interview of a woman (Figure 1). Further, to be eligible to be administered both the list and

[5] The specific question is worded as follows: Please imagine a ten-step ladder, where on the bottom, the first step, are those who are totally unable to change their lives, while on step 10, the highest step, stand those who have full control over their own life. On which step are you? Respondents specify a value between one and ten.

[6] There was also a separate coding error in which a subset of these 45% were not in fact administered the direct response module, but this error is not relevant to the list experiment analysis conducted here.

direct response IPV questions, women had to be living with their husbands/partners in the last 13 months and be alone or with a child less than 36 months at the time of the interview. Among the 2,083 list respondents, 1757 (or 84 percent) were living with their husbands/partners and a further 1,418 (or 68 percent) were also alone or with a child less than 36 months at the time of the survey. Lastly, 19 women refused to be administered the direct response questions; thus the sample for analysis includes 1399 women who were administered both the direct response and list questions.

We analyze whether there is differential attrition across the four list randomization groups (A through D), which may bias the results . In the sample assigned to receive the control list (group list D), 32% of respondents attrite from the list experiment for one of the reasons noted above (absence of cohabitation, failure to be alone at the point of the survey, or absence of consent). The observed attrition rate is slightly higher in the list treatment groups and particularly in list group C, but the differences are not statistically significant.[7]

3. Estimation/Methods

Our primary objective in the empirical analysis is to estimate the prevalence of various forms of IPV using the list experiment and compare this prevalence to the corresponding prevalence identified using direct questions. In addition, we analyze the difference in prevalence between the direct response and list experiment among various subsamples of interest.

For the primary analysis of the list experiment, we regress the number of list items that the respondent concurs with on a binary variable for assignment to the list treatment arm, as well as woreda fixed effects given that randomization was stratified at the woreda level; standard errors are clustered at the kebele level. The resulting coefficient and standard error (β_{LE}) from the treatment indicator correspond to the estimated prevalence of IPV generated using the list experiment. We report this estimate in conjunction with the estimated prevalence using a direct question (β_{DR}) and the difference between the two; we also report the p-value corresponding to the test $\beta_{DR} = \beta_{LE}$. To assess whether there is any systematic difference between the prevalence rates estimated using direct reports and the list experiments, we report the p-value corresponding to a chi squared test of the joint hypothesis that $\beta_{DR} = \beta_{LE}$ across all nine items, as well as for each domain of violence constituted by three items (emotional violence, physical violence and sexual violence). We then explore similar patterns in various subsamples of interest.

The validity of prevalence estimates derived from list randomization depends on a set of assumptions that have been developed in previous literature (particularly Blair and Imai, 2013). In particular, the required assumptions include no design effect (the inclusion of the sensitive item does not shift respondents' reaction to the non-sensitive items) and no floor or ceiling effects (in which responses are uniformly negative or affirmative to the list items). We report in Table A1 the frequency of responses for each numerical category for both control and treatment respondents for each list, noting that there

[7] The probability of attrition is regressed on the binary variables corresponding to list group assignment with list group D as the omitted group; standard errors are clustered at the kebele level. The coefficient is 0.018 for list group A (p=0.533), 0.014 for list group B (p=0.669), and 0.050 for list group C (p=0.104).

is little evidence of ceiling and floor effects. In addition, we conduct a formal test for the presence of design effects, and fail to reject the null hypothesis of no design effects for each of the three lists.[8]

4. Results

a. Summary statistics

Table 1 reports summary statistics for the sample in each of the four list randomization arms (respondents randomized to list groups A, B, C, and control). The final column of the table reports a balance test where the covariates of interest are regressed on binary variables for assignment to list groups A, B, and C, and we report the p-value on the test $\beta_{ListA} = \beta_{ListB} = \beta_{ListC}$. This test corresponds to the null hypothesis for the joint test that all four subgroups (the control group, and list groups A, B, and C) are comparable on observable characteristics.

On average, female respondents are around 31 years old, and approximately a quarter report some education; their spouses are around 39 years old, and 40% report some education. 10% of households are female-headed households, and the sample is overwhelmingly primarily engaged in crop cultivation.

Reported rates of IPV are relatively low in this sample relative to nationwide prevalence rates in Ethiopia (e.g., derived from the Demographic and Health Survey): 11% of the sample reports past-year experience of emotional IPV, 8% reports past-year experience of physical IPV, and 4% reports past-year experience of sexual IPV. In order to capture norms around violence against women, both female and male respondents are also posed a series of questions about whether IPV is justified under certain circumstances; 43% of women and 65% of men state that IPV is not justified under any of the enumerated circumstances.[9]

In general, the evidence suggests the sample is balanced across the list randomization groups with respect to demographic characteristics. However, the rate of physical IPV and thus physical and/or sexual IPV is significantly higher in list group A. As reported in the final row, a joint test of balance across all covariates examined fails to reject the hypothesis that covariates are comparable across arms (p=.946).

b. LIST method versus Direct Method

i. Primary results

The primary results analyzing the list experiment vis-à-vis the direct report questions are reported in Table 2. As noted in Table 1, in general the rate of directly reported past-year experience of IPV is relatively low, ranging between 3% and 8%. For estimated prevalence measured using the list experiment, we observe prevalence rates that are close to zero or, when positive, in the same range as the direct report estimates. We cannot reject the hypothesis that $\beta_{DR} = \beta_{LE}$ for any of the nine forms of IPV measured, suggesting that the list experiment is not differentially eliciting reports of IPV vis-à-vis the direct report questions. In addition, there is no meaningful difference in these results when estimated

[8] For list 1 as enumerated in Appendix 1, the p-value associated with this test is .364; for list 2 , the p-value is .626; for list 3 , the p-value is .518.

[9] The circumstances include if the wife goes out without telling her husband; if she neglects the children; if she argues with her husband; if she burns food; and if she refuses to have sex with her husband.

conditional on a set of basic demographic controls (the education level of the household head, the education level of the primary female, the age of both the primary male and the primary female, the number of children, region fixed effects, sex of the enumerator, and binary variables for treatment arm).[10]

In addition, four of the prevalence rates estimated using the list experiment are negative, suggesting that on average, respondents who were presented with the "treatment" list including the sensitive, IPV-related item concurred with fewer items than respondents who were presented with the "control" list including only the baseline items. If we compare the estimated prevalence rates drawn from the list experiment for different types of violence, in general the prevalence rates for emotional IPV and physical IPV are positive for the less severe forms of violence and negative for the more severe forms (Threatened to hurt or harm you or someone you care about; twisted your arm or pulled your hair). Larger and more consistent negative point estimates are observed for sexual IPV. This pattern suggests that the risk of negative prevalence rates is particularly high for more sensitive items such as severe emotional and physical IPV and sexual IPV. This is an unusual pattern in the list experiment data that is examined further in the next subsection.

ii. Explaining fleeing behavior

One hypothesis for the unusual pattern observed in responses to list experiments is that respondents are "fleeing" from the sensitive item in the treatment lists. The documentation of fleeing behavior (as observed in negative point estimates) is relatively novel in the literature. Chuang et al. (2021) report in their analysis of a double list experiment in Cote d'Ivoire some evidence of negative point estimates generated from list experiments and argue that respondents are reluctant to provide truthful answers even for non-sensitive items, given a fear that researchers will infer information about their behavior linked to sensitive items. They also find that respondents are unwilling to report sensitive behavior even when anonymity is guaranteed through random response techniques and posit that this is due to self-protective behavior, especially among respondents who behave according to the social norm.

We further probe this fleeing hypothesis by investigating heterogeneity with respect to respondent, enumerator and list characteristics, in order to identify subsamples of respondents and situations in which fleeing behavior might be more pronounced. In particular, we examine subsamples in which IPV might be more stigmatized and respondents less likely to admit being victims, and where respondents trust less the anonymity of the list experiment; these include subsamples defined by age, education, self-efficacy, whether the respondent states that IPV is justified, and sex of the enumerator. We analyze these subsamples for the different types of violence and for the different set of lists.

We summarize the heterogeneity analysis in Table 3. For each subsample, we report the mean difference (β_{DR} - β_{LE}) for each form of violence (emotional, physical and sexual IPV) in conjunction with the p-value corresponding to the test of the hypothesis that the two estimates are equal for this form of violence. We also report the analogous mean difference and p-value across all types of violence. Note that a negative mean difference suggests that on average, more reports of IPV are elicited through the

[10] These results are reported in Appendix Table A2.

list experiment vis-à-vis direct reports, while a positive mean difference suggests more reports of IPV are elicited through direct reports (consistent with what we are describing as fleeing behavior). The detailed results for each subsample are reported in Tables A3 through A7 in the Appendix.

Table 3 reveals no evidence of fleeing behavior across subgroups that have some or no education or that concur or do not concur that IPV is justified. There is weak evidence of heterogeneity by age and by enumerator sex; fleeing behavior is concentrated among respondents who are younger than 30 and among respondents who are surveyed by a male enumerator, but the differences are not statistically significant. By contrast, when we examine the results for women with higher and lower reported self-efficacy, we observe a meaningfully different pattern. Within the sample of women characterized by low self-efficacy (below the median), we observe negative mean differences in IPV prevalence – i.e., these women are more likely to report IPV using the list experiment – and the difference is statistically significant for physical IPV (p=0.023). Within the sample of women characterized by high self-efficacy (equal to or above the median), we observe positive mean differences in IPV prevalence – i.e., these women exhibit fleeing behavior – and the differences are statistically significant at the ten percent level for emotional IPV (p=0.091), physical IPV (p=0.068), and pooled IPV (p=0.062).

In Table 4, we report a parallel analysis by the control list employed, rather than the type of IPV elicited. As noted above, two of the three control lists (list one and list two) included other sensitive items, while list three did not include any sensitive items. Here, there is some evidence that fleeing behavior is most pronounced for list three; the difference between the direct report and list experiment prevalence estimates is significant when using this list for two subsamples, women who do not concur that IPV is justified (p=0.076) and women characterized by high self-efficacy (p=0.023).

To sum up, we find that respondent and control list characteristics are meaningful determinants of fleeing behavior. Respondents who are characterized by high self-efficacy show more fleeing behavior, and this is also more pronounced for control lists that do not include other sensitive items (in which case fleeing behavior is also observed for women who concur that violence is not justified). One potential hypothesis is that women who are characterized by higher self-efficacy and decline to justify violence are particularly reluctant to be associated with a list including IPV and so exhibit fleeing behavior. More specifically, women who seek to report non-experience of violence can simply report zero in a direct question (where zero constitutes a lower bound); when posed a list question, however, they can exhibit fleeing behavior and drive the estimated prevalence down below zero. Appendix Table A5 reveals that for women with higher self-efficacy, questions where women exhibit fleeing behavior also report very low levels of violence, consistent with the hypothesis that responses among women who report no experience of IPV is driving the observed pattern. Alternatively, these respondents may be less attentive to the list question.

5. Conclusion

In this paper, we document that a list experiment conducted in rural Ethiopia to measure the prevalence of emotional, physical and sexual IPV was not effective in eliciting higher reports of violence vis-à-vis direct questions administered to the same sample. The estimated prevalence rates are not significantly different across the two methods, and in a number of cases the estimated prevalence constructed using the list method is in fact lower (or even negative). We present suggestive evidence that this may reflect

"fleeing" behavior on the part of respondents, who seek to distance themselves from a list that includes items linked to IPV. This seems to be particularly true for more severe/sensitive forms of violence (severe emotional and physical violence and sexual violence). We further probe this fleeing hypothesis by investigating heterogeneity with respect to respondent, enumerator and list characteristics and find that respondents who are characterized by high self-efficacy show more fleeing behavior, and this is also more pronounced for control lists that do not include other sensitive items.

Our results emphasize the importance of proceeding with caution when utilizing list experiments, particularly in samples where IPV is highly stigmatized or levels of trust are low. List experiments can easily fail for various reasons leading to prevalence rates that are extremely hard to interpret. Identifying other effective methods to measure sensitive health behaviors remains an important direction for future research.

Bibliography

Agüero, J. M., & Frisancho, V. (2021). Measuring violence Against Women with Experimental Methods. *Economic Development and Cultural Change*. https://doi.org/10.1086/714008

Bell, S. O., & Bishai, D. (2019). Can a List Experiment Improve Validity of Abortion Measurement? *Studies in Family Planning*, *50*(1), 43–61. https://doi.org/10.1111/sifp.12082

Blair, G., Coppock, A., & Moor, M. (2020). When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments. *American Political Science Review*. https://doi.org/10.1017/S0003055420000374

Blair, G., & Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis*. https://doi.org/10.1093/pan/mpr048

Bulte, E., & Lensink, R. (2019). Women's empowerment and domestic abuse: Experimental evidence from Vietnam. *European Economic Review*, *115*, 172–191. https://doi.org/10.1016/j.euroecorev.2019.03.003

Campbell, J. C. (2002). Health consequences of intimate partner violence. *Lancet*, *359*(9314), 1331–1336. https://doi.org/10.1016/S0140-6736(02)08336-8

Castilla, C., & Murphy, D. (2021). *Identifying Intimate Partner Violence by Gender: Evidence from a List Experiment in Kenya*. https://doi.org/dx.doi.org/10.2139/ssrn.3771604

Chong, A., Gonzalez-Navarro, M., Karlan, D., & Valdivia, M. (2020). Do Information Technologies Improve Teenagers' Sexual Education? Evidence from a Randomized Evaluation in Colombia. *World Bank Economic Review*, *34*(2), 371–392. https://doi.org/10.1093/wber/lhy031

Chuang, E., Dupas, P., Huillery, E., & Seban, J. (2021). Sex, lies, and measurement: Consistency tests for indirect response survey methods. *Journal of Development Economics*. https://doi.org/10.1016/j.jdeveco.2020.102582

Cullen, C. A. (2020). Method Matters : Underreporting of Intimate Partner Violence in Nigeria and Rwanda. *World Bank Policy Research Working Paper*, *9274*, 1–40.

De Cao, E., & Lutz, C. (2018). Sensitive Survey Questions: Measuring Attitudes Regarding Female Genital Cutting Through a List Experiment. *Oxford Bulletin of Economics and Statistics*, *80*(5), 871–892. https://doi.org/10.1111/obes.12228

Devries, K. M., Mak, J. Y., Bacchus, L. J., Child, J. C., Falder, G., Petzold, M., Astbury, J., & Watts, C. H. (2013). Intimate Partner Violence and Incident Depressive Symptoms and Suicide Attempts: A Systematic Review of Longitudinal Studies. *PLoS Medicine*, *10*(5). https://doi.org/10.1371/journal.pmed.1001439

Ellsberg, M., Jansen, H. A., Heise, L., Watts, C. H., & Garcia-Moreno, C. (2008). Intimate partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. *The Lancet*, *371*(9619), 1165–1172.

Haber, N., Harling, G., Cohen, J., Mutevedzi, T., Tanser, F., Gareta, D., Herbst, K., Pillay, D., Bärnighausen, T., & Fink, G. (2018). List randomization for eliciting HIV status and sexual behaviors in rural KwaZulu-Natal, South Africa: A randomized experiment using known true values for validation. *BMC Medical Research Methodology*, *18*(1), 1–12. https://doi.org/10.1186/s12874-018-0507-9

Hinsley, A., Keane, A., St. John, F. A. V., Ibbett, H., & Nuno, A. (2019). Asking sensitive questions using the unmatched count technique: Applications and guidelines for conservation. *Methods in Ecology and Evolution*, *10*(3), 308–319. https://doi.org/10.1111/2041-210X.13137

Joseph, G., Javaid, S. U., Andres, L. A., Chellaraj, G., Solotaroff, J. L., & Rajan, S. I. (2017). Underreporting of Gender-Based Violence in Kerala, India: An Application of the List Randomization Method. *Underreporting of Gender-Based Violence in Kerala, India: An Application of the List Randomization Method*, *April*. https://doi.org/10.1596/1813-9450-8044

Kotsadam, A., & Villanger, E. (2019). Jobs and Intimate Partner Violence - Evidence from a Field Experiment in Ethiopia. *Working Paper*, *February*.

Lépine, A., Treibich, C., & D'Exelle, B. (2020). Nothing but the truth: Consistency and efficiency of the list experiment method for the measurement of sensitive health behaviours. *Social Science and Medicine*, *266*, 113326. https://doi.org/10.1016/j.socscimed.2020.113326

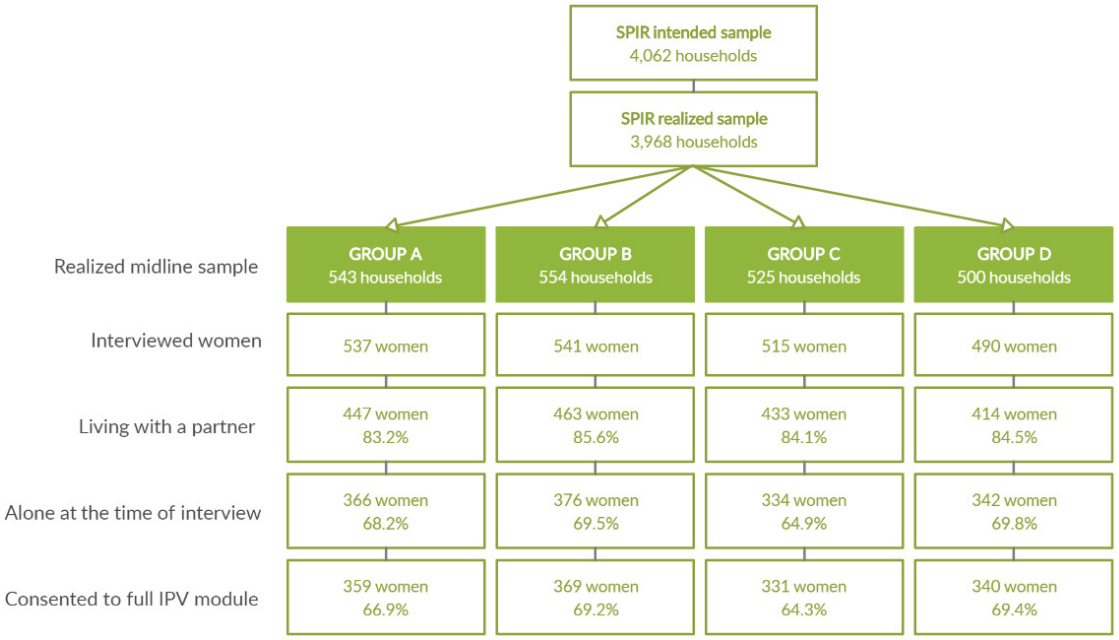
Moseson, H., Massaquoi, M., Dehlendorf, C., Bawo, L., Dahn, B., Zolia, Y., Vittinghoff, E., Hiatt, R. A., & Gerdts, C. (2015). Reducing under-reporting of stigmatized health events using the List Experiment: Results from a randomized, population-based study of abortion in Liberia. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyv174

Moseson, H., Treleaven, E., Gerdts, C., & Diamond-Smith, N. (2017). The List Experiment for Measuring Abortion: What We Know and What We Need. In *Studies in Family Planning*. https://doi.org/10.1111/sifp.12042

Traunmuller, R., Kijewski, S., & Freitag, M. (2018). The Silent Victims of Wartime Sexual Violence: Evidence from a List Experiment in Sri Lanka. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2937943

Treibich, C., & Lépine, A. (2019). Estimating misreporting in condom use and its determinants among sex workers: Evidence from the list randomisation method. *Health Economics (United Kingdom)*, *28*(1), 144–160. https://doi.org/10.1002/hec.3835

Figure 1: Flow chart of sample of women randomized into LIST groups.



Notes: This figure presents a summary of the sample for analysis of the list experiment by treatment arm. The percentages calculate the percent of the realized sample reported in each cell.

Table 1: Summary statistics (mean or percentage) and balance

	List A	N	List B	N	List C	N	List D (control)	N	p-value
Age of primary female	30.955	359	31.06	366	31.424	330	30.826	340	0.754
Primary female has some education	0.258	357	0.269	361	0.264	326	0.286	339	0.858
Age of primary male	39.155	349	39.209	358	39.183	323	38.827	329	0.946
Primary male has some education	0.416	346	0.398	357	0.425	322	0.473	328	0.189
Household size	6.259	359	6.092	369	6.311	331	6.232	340	0.464
Female-headed household	0.095	359	0.089	369	0.073	330	0.103	340	0.616
Household head's main activity is crop production	0.896	356	0.907	365	0.921	329	0.891	340	0.402
Experienced emotional violence in the past 13 months	0.125	359	0.089	369	0.121	331	0.115	340	0.291
Experienced physical violence in the past 13 months	0.109	359	0.051	369	0.076	331	0.068	340	0.028
Experienced sexual violence in the past 13 months	0.033	359	0.024	369	0.057	331	0.038	340	0.129
Experienced physical or sexual violence in the past 13 months	0.125	359	0.07	369	0.106	331	0.082	340	0.037
A husband is not justified in beating his wife in any of these situations (female response)	0.446	354	0.441	365	0.41	329	0.419	339	0.677
A husband is not justified in beating his wife in any of these situations (male response)	0.682	321	0.633	338	0.624	298	0.655	296	0.326
P-value: Joint test									0.946

Note: P-value in the last column reports the results of an F-test of equality of means for Lists A, B, C and control. The final p-value is a test of joint significance across all characteristics.

Table 2: Analysis of list experiment

	Direct report	N	List report	N	Difference	p-value
Emotional violence: Said something to humiliate you in front of others	0.041	1399	0.018	699	0.022	0.782
	(0.005)		(0.081)			
Emotional violence: Threatened to hurt or harm you or someone you care about	0.02	1399	-0.037	698	0.057	(0.466)
	(0.004)		(0.079)			
Emotional violence: Insulted you or made you feel bad about yourself	0.103	1399	0.045	698	0.057	(0.493)
	(0.009)		(0.083)			
Physical violence: Pushed you, shook you, or thrown something at you	0.038	1399	0.067	709	-0.029	(0.725)
	(0.005)		(0.082)			
Physical violence: Slapped you	0.054	1399	0.013	709	0.04	(0.613)
	(0.007)		(0.080)			
Physical violence: Twisted your arm or pulled your hair	0.026	1399	-0.031	708	0.057	(0.493)
	(0.005)		(0.069)			
Sexual violence: Physically forced you to have sexual intercourse	0.034	1399	0.008	670	0.026	(0.765)
	(0.005)		(0.088)			
Sexual violence: Physically forced you to perform any other sexual acts	0.022	1397	-0.026	669	0.049	(0.563)
	(0.004)		(0.085)			
Sexual violence: Forced you with threats/in another way to perform sexual acts	0.016	1398	-0.065	670	0.082	(0.283)
	(0.004)		(0.077)			
Joint (chi-squared) test across all nine items						0.363
Joint test: emotional violence items						0.427
Joint test: Physical violence items						0.524
Joint test: Sexual violence items						0.397

Notes: The direct report column reports the prevalence estimate for the specified act of violence in the last 13 months derived from the direct report questions; the list experiment column reports the prevalence estimate for the specified act derived from an analysis of the list experiment question, in which the respondent's answer is regressed on a binary variable for treatment list, conditional on woreda fixed effects and with standard errors clustered at the kebele level. The final column reports the p-value for a test of equality across the direct report and list experiment estimates. The final row reports a joint test of equality across the direct report and list experiment questions, pooling across all nine items.

Table 3: Analysis of list experiment by subsamples and type of violence

		Emotional violence		**Physical violence**		**Sexual violence**		**Pooled violence**	
	Obs.	Average difference	p-value (joint test)	Average difference	p-value (joint test)	Average difference	p-value (joint test)	Average difference	p-value (joint test)
Age									
Under 30	1501	0.063	0.474	0.076	0.416	0.146	0.112	0.095	0.197
30 or older	2534	0.038	0.643	-0.018	0.824	-0.01	0.899	0.003	0.956
Education									
Any education	982	0.063	0.568	0.079	0.482	-0.088	0.416	0.018	0.835
No education	2843	0.03	0.672	-0.009	0.901	0.11	0.147	0.044	0.448
Self-efficacy									
Higher self-efficacy	2365	0.134	0.091	0.147	0.068	0.07	0.367	0.117	0.062
Lower self-efficacy	1441	-0.128	0.243	-0.219	0.023	-0.028	0.796	-0.125	0.132
Violence perceptions									
Violence is never justified	1595	0.109	0.283	0.015	0.875	0.035	0.746	0.053	0.51
Violence is justified	2228	0.01	0.907	0.015	0.851	0.058	0.475	0.027	0.684
Sex of Enumerator									
Female enumerator	2154	-0.003	0.972	0	0.996	-0.044	0.638	-0.016	0.83
Male enumerator	1098	0.07	0.38	0.015	0.862	0.076	0.373	0.054	0.425

Table 4: Analysis of list experiment by subsamples and control list type

		Control list 1		**Control list 2**		**Control list 3**	
	Obs.	Average difference	p-value (joint test)	Average difference	p-value (joint test)	Average difference	p-value (joint test)
Age							
Under 30	1501	0.036	0.735	0.145	0.134	0.104	0.272
30 or older	2534	-0.01	0.899	-0.02	0.806	0.041	0.628
Education							
Any education	982	0.03	0.817	-0.03	0.808	0.054	0.643
No education	2843	-0.008	0.913	0.078	0.286	0.061	0.394
Self-efficacy							
Higher self-efficacy	2365	0.032	0.689	0.136	0.117	0.184	0.023
Lower self-efficacy	1441	-0.115	0.287	-0.12	0.261	-0.14	0.128
Violence perceptions							
Violence is never justified	1595	-0.011	0.919	-0.004	0.969	0.174	0.076
Violence is justified	2228	0.028	0.755	0.092	0.294	-0.038	0.631
Enumerator gender							
Female enumerator	2154	-0.147	0.144	0.096	0.304	0.003	0.974
Male enumerator	1098	0.106	0.219	-0.005	0.954	0.061	0.431

Appendix 1: LIST questions

Instructions for CAPI:

- At the household level, randomly assign 50% of households to receive the LIST module and 50% of households not to receive the LIST module
- For the 50% of households that receive the LIST module, randomly assign them into 1 of 4 groups (A, B, C, D)
- Mothers in Group A, B,C,D will receive the 3 lists below. But question 5 in each list will vary depending on what group the mother is in. So for example, a mother randomly assigned to **Group A** will receive LIST 1, LIST 2, and LIST 3, and within each list she will receive question **5a**. A mother randomly assigned to **Group B** will receive all three lists, and within each list she will receive question **5b**, etc…

(LIST 1) I5_11. Here is a list of five (CAPI replace with "four" if randomized to group D) things that some people have done or experienced, and some people have not. Please listen to them and then tell me HOW MANY of them you have done or experienced in the last 13 months. Do not tell me which you have experienced, just tell me how many total. Wait until I have read the entire list to respond. I will read the entire list 2 times. If you need me to read it again, please let me know. In the past 13 months: **[**CAPI, place 5a, 5b, or 5c in the third place**]**	I5_11. RESPONSE _____

1	Have you taken care of a sick relative who is unable to care for themselves
2	Have you heard of a girl who has been forced to marry a man against her will or wishes
3	Has your husband traveled to the kebele/village center to buy goods or sell items produced by your household
4	Have you and your husband/spouse attended the wedding celebration of a friend
Randomly administer one of the following:	
5a	Has your husband/spouse said or done something to humiliate you in front of others
5b	Has your husband/spouse pushed you, shook you, or thrown something at you?
5c	Has your husband/spouse physically forced you to have sexual intercourse with him when you did not want to
5d	BLANK

(LIST 2) I5_12. Again, here is a list of five (CAPI replace with "four" if randomized to group D) things that some people have done or experienced, and some people have not. Please listen to them and then tell me HOW MANY of them you have done or experienced in the last 13 months. Do not tell me which you have experienced, just tell me how many total. Wait until I have read the entire list to respond. I will read the entire list 2 times. If you need me to read it again, please let me know. In the past 13 months: **[**CAPI, place 5a, 5b, or 5c in the second place**]**	I5_12. RESPONSE _____

1	Have you and your husband attended a funeral of a friend or family member
2	Had a relative or close family friend's daughter undergo genital circumcision
3	Has your husband/spouse traveled to a market to buy livestock (small or large)
4	Have you traveled outside your community to visit a friend
Randomly administer one of the following:	
5a	Has your husband/spouse threatened to hurt or harm you or someone you care about
5b	Has your husband/spouse slapped you?
5c	Has your husband/spouse physically forced you to perform any other sexual acts you did not want to?
5d	BLANK

(LIST 3) I5_13 Again, here is a list of five (CAPI replace with "four" if randomized to group D) things that some people have done or experienced, and some people have not. Please listen to them and then tell me HOW MANY of them you have done or experienced in the last 13	I5_13. RESPONSE

months. Do not tell me which you have experienced, just tell me how many total. Wait until I have read the entire list to respond. I will read the entire list 2 times. If you need me to read it again, please let me know. In the past 13 months: **[**CAPI, place 5a, 5b, or 5c in the fourth place**]**		_____
1	Has your husband/spouse became too sick to work or attend to daily chores	
2	Have you used a mobile phone to make a call to a friend	
3	Have you traveled to a market to purchase clothes for yourself and/or child	
4	Have you and your husband/spouse had any money or items stolen from you	
Randomly administer one of the following		
5a	Has your husband/spouse insulted you or made you feel bad about yourself?	
5b	Has your husband/spouse twisted your arm or pulled your hair	
5c	Has your husband/spouse force you with threats or in any other way to perform sexual acts you did not want to?	
5d	BLANK	

Appendix 2: Additional tables

Table A1: Distribution of responses in the list randomization experiment

List 1	Control set (excluding IPV item)	Treatment set (including IPV item)
0	0.225	0.260
1	0.339	0.272
2	0.231	0.273
3	0.164	0.143
4	0.041	0.033
5	0.000	0.019
List 2	Control	Treatment
0	0.278	0.305
1	0.330	0.297
2	0.249	0.253
3	0.108	0.111
4	0.035	0.024
5	0.000	0.009
List 3	Control	Treatment
0	0.249	0.248
1	0.330	0.364
2	0.284	0.254
3	0.120	0.107
4	0.018	0.021
5	0.000	0.007

Table A2: Analysis of list experiment conditional on demographic control variables

	Direct report	N	List report	N	Difference	p-value
Emotional violence: Said something to humiliate you in front of others	0.041	1399	0.023	673	0.018	(0.827)
	(0.005)		(0.083)			
Emotional violence: Threatened to hurt or harm you or someone you care about	0.02	1399	-0.015	672	0.035	(0.642)
	(0.004)		(0.076)			
Emotional violence: Insulted you or made you feel bad about yourself	0.103	1399	0.041	672	0.062	(0.442)
	(0.009)		(0.080)			
Physical violence: Pushed you, shook you, or thrown something at you	0.038	1399	0.055	678	-0.017	(0.830)
	(0.005)		(0.081)			
Physical violence: Slapped you	0.054	1399	0.045	678	0.009	(0.911)
	(0.007)		(0.081)			
Physical violence: Twisted your arm or pulled your hair	0.026	1399	-0.017	677	0.062	(0.442)
	(0.005)		(0.069)			
Sexual violence: Physically forced you to have sexual intercourse	0.034	1399	0.012	644	0.021	(0.803)
	(0.005)		(0.087)			
Sexual violence: Physically forced you to perform any other sexual acts	0.022	1397	-0.006	643	0.029	(0.735)
	(0.004)		(0.086)			
Sexual violence: Forced you with threats/in another way to perform sexual acts	0.016	1398	-0.071	644	0.087	(0.251)
	(0.004)		(0.077)			
Joint (chi-squared) test across all nine items						0.321
Joint test: emotional violence items						0.414
Joint test: Physical violence items						0.849
Joint test: Sexual violence items						0.331

Notes: The direct report column reports the prevalence estimate for the specified act of violence in the last 13 months derived from the direct report questions; the list experiment column reports the prevalence estimated for the specified act derived from an analysis of the list experiment question, in which the respondent's answer is regressed on a binary variable for treatment list, conditional on a set of demographic control variables (the education level of the household head, the education level of the primary female, the age of both the primary male and the primary female, the number of children, region fixed effects, enumerator sex and binary variables for treatment arm) and with standard errors clustered at the kebele level. The final column reports the p-value for a test of equality across the direct report and list experiment estimates.

Table A3: Subsamples of younger (below age 30) and older (above 30) women

	Age < 30				Age ≥ 30			
	Direct report	**List report**	**Difference**	**p-value**	**Direct report**	**List report**	**Difference**	**p-value**
Emotional violence: Said something to humiliate you	0.035	0.037	-0.003	(0.984)	0.045	0.023	0.021	(0.850)
	(0.008)	(0.134)			(0.007)	(0.115)		
Emotional violence: Threatened to hurt or harm you	0.017	-0.117	0.135	(0.232)	0.022	0.041	-0.019	(0.863)
	(0.005)	(0.112)			(0.005)	(0.111)		
Emotional violence: Insulted you or made you feel bad about yourself	0.099	0.112	-0.013	(0.919)	0.105	0.012	0.093	(0.409)
	(0.013)	(0.128)			(0.011)	(0.111)		
Physical violence: Pushed you, shook you, or thrown something	0.035	-0.027	0.062	(0.643)	0.04	0.076	-0.036	(0.763)
	(0.007)	(0.133)			(0.007)	(0.120)		
Physical violence: Slapped you	0.059	-0.066	0.126	(0.339)	0.05	0.065	-0.015	(0.895)
	(0.010)	(0.131)			(0.007)	(0.113)		
Physical violence: Twisted your arm or pulled your hair	0.021	-0.029	0.05	(0.691)	0.029	-0.05	0.079	(0.443)
	(0.006)	(0.127)			(0.006)	(0.102)		
Sexual violence: Physically forced you to have sexual intercourse	0.03	-0.043	0.072	(0.619)	0.036	0.027	0.009	(0.936)
	(0.007)	(0.146)			(0.007)	(0.113)		
Sexual violence: Physically forced you to perform any other sexual act	0.021	-0.093	0.113	(0.389)	0.023	0.057	-0.034	(0.742)
	(0.006)	(0.133)			(0.006)	(0.102)		
Sexual violence: Forced you with threats/in another way to perform sexual acts	0.019	-0.178	0.197	(0.094)	0.015	0.018	-0.004	(0.971)
	(0.006)	(0.119)			(0.004)	(0.099)		

Notes: The direct report column reports the prevalence estimate derived from the direct report questions of violence in the last 13 months; the list experiment column reports the prevalence estimate derived from an analysis of the list experiment question, in which the respondent's answer is regressed on a binary variable for treatment list, conditional on woreda fixed effects and with standard errors clustered at the kebele level. The final column reports the p-value for a test of equality across the direct report and list experiment estimates.

Table A4: Subsamples of women who do and do not report any education

	No education				Some education			
	Direct report	**List report**	**Difference**	**p-value**	**Direct report**	**List report**	**Difference**	**p-value**
Emotional violence: Said something to humiliate you	0.046	0.071	-0.025	(0.769)	0.024	-0.068	0.093	(0.586)
	(0.007)	(0.086)			(0.008)	(0.175)		
Emotional violence: Threatened to hurt or harm you	0.023	-0.014	0.037	(0.694)	0.013	-0.107	0.121	(0.422)
	(0.005)	(0.094)			(0.006)	(0.155)		
Emotional violence: Insulted you or made you feel bad about yourself	0.112	0.034	0.078	(0.403)	0.083	0.107	-0.024	(0.882)
	(0.010)	(0.092)			(0.014)	(0.164)		
Physical violence: Pushed you, shook you, or thrown something	0.039	0.136	-0.098	(0.313)	0.035	-0.096	0.131	(0.390)
	(0.006)	(0.098)			(0.010)	(0.156)		
Physical violence: Slapped you	0.05	0.011	0.04	(0.651)	0.062	0.047	0.015	(0.924)
	(0.007)	(0.088)			(0.013)	(0.159)		
Physical violence: Twisted your arm or pulled your hair	0.026	-0.007	0.032	(0.691)	0.024	-0.066	0.09	(0.545)
	(0.005)	(0.082)			(0.008)	(0.154)		
Sexual violence: Physically forced you to have sexual intercourse	0.041	-0.057	0.098	(0.324)	0.016	0.151	-0.135	(0.406)
	(0.006)	(0.101)			(0.007)	(0.170)		
Sexual violence: Physically forced you to perform any other sexual act	0.027	-0.131	0.158	(0.084)	0.011	0.235	-0.225	(0.138)
	(0.005)	(0.092)			(0.005)	(0.158)		
Sexual violence: Forced you with threats/in another way to perform sexual acts	0.02	-0.054	0.074	(0.420)	0.008	-0.087	0.095	(0.469)
	(0.004)	(0.092)			(0.005)	(0.137)		

Notes: The direct report column reports the prevalence estimate derived from the direct report questions of violence in the last 13 months; the list experiment column reports the prevalence estimate derived from an analysis of the list experiment question, in which the respondent's answer is regressed on a binary variable for treatment list, conditional on woreda fixed effects and with standard errors clustered at the kebele level. The final column reports the p-value for a test of equality across the direct report and list experiment estimates.

Table A5: Subsamples of women characterized by relatively lower and higher self-efficacy

	Lower self-efficacy				Higher self-efficacy			
	Direct report	**List report**	**Difference**	**p-value**	**Direct report**	**List report**	**Difference**	**p-value**
Emotional violence: Said something to humiliate you	0.063	0.153	-0.09	(0.510)	0.028	-0.037	0.065	(0.508)
	(0.011)	(0.141)			(0.006)	(0.100)		
Emotional violence: Threatened to hurt or harm you	0.028	0.177	-0.15	(0.292)	0.016	-0.138	0.153	(0.122)
	(0.007)	(0.146)			(0.004)	(0.100)		
Emotional violence: Insulted you or made you feel bad about yourself	0.117	0.261	-0.144	(0.269)	0.095	-0.059	0.154	(0.159)
	(0.014)	(0.133)			(0.010)	(0.110)		
Physical violence: Pushed you, shook you, or thrown something	0.043	0.301	-0.258	(0.046)	0.035	-0.038	0.072	(0.477)
	(0.009)	(0.133)			(0.006)	(0.104)		
Physical violence: Slapped you	0.063	0.24	-0.177	(0.156)	0.048	-0.086	0.134	(0.214)
	(0.011)	(0.128)			(0.007)	(0.109)		
Physical violence: Twisted your arm or pulled your hair	0.03	0.251	-0.222	(0.042)	0.024	-0.203	0.227	(0.011)
	(0.008)	(0.112)			(0.005)	(0.091)		
Sexual violence: Physically forced you to have sexual intercourse	0.045	0.044	0.001	(0.993)	0.027	0.023	0.003	(0.973)
	(0.009)	(0.144)			(0.005)	(0.104)		
Sexual violence: Physically forced you to perform any other sexual act	0.032	0.064	-0.033	(0.806)	0.017	-0.076	0.093	(0.379)
	(0.008)	(0.136)			(0.004)	(0.107)		
Sexual violence: Forced you with threats/in another way to perform sexual acts	0.024	0.077	-0.053	(0.647)	0.012	-0.157	0.17	(0.063)
	(0.007)	(0.119)			(0.004)	(0.092)		

Notes: The direct report column reports the prevalence estimate derived from the direct report questions of violence in the last 13 months; the list experiment column reports the prevalence estimate derived from an analysis of the list experiment question, in which the respondent's answer is regressed on a binary variable for treatment list, conditional on woreda fixed effects and with standard errors clustered at the kebele level. The final column reports the p-value for a test of equality across the direct report and list experiment estimates. Lower self-efficacy indicates respondents who are below the median response while higher self-efficacy indicates respondents who are at or above the median response.

Table A6: Subsample of respondents who do and do not concur with justifications for IPV

	Panel A: IPV is unjustified sample				Panel B: IPV is justified sample			
	Direct report	**List report**	**Difference**	**p-value**	**Direct report**	**List report**	**Difference**	**p-value**
Emotional violence: Said something to humiliate you	0.03	-0.045	0.076	(0.555)	0.049	0.031	0.019	(0.857)
	(0.007)	(0.131)			(0.008)	(0.106)		
Emotional violence: Threatened to hurt or harm you or others	0.015	-0.021	0.036	(0.780)	0.024	-0.064	0.088	(0.403)
	(0.005)	(0.133)			(0.005)	(0.106)		
Emotional violence: Insulted you or made you feel bad about yourself	0.074	-0.142	0.216	(0.097)	0.124	0.201	-0.077	(0.483)
	(0.011)	(0.134)			(0.012)	(0.109)		
Physical violence: Pushed you, shook you, or thrown something	0.029	0.103	-0.074	(0.557)	0.046	0.053	-0.008	(0.943)
	(0.007)	(0.128)			(0.007)	(0.109)		
Physical violence: Slapped you	0.047	0.063	-0.016	(0.895)	0.059	-0.025	0.084	(0.417)
	(0.009)	(0.124)			(0.008)	(0.105)		
Physical violence: Twisted your arm or pulled your hair	0.02	-0.114	0.134	(0.212)	0.03	0.062	-0.032	(0.724)
	(0.006)	(0.110)			(0.006)	(0.090)		
Sexual violence: Physically forced you to have sexual intercourse	0.027	0.06	-0.033	(0.810)	0.039	-0.033	0.073	(0.516)
	(0.007)	(0.143)			(0.007)	(0.115)		
Sexual violence: Physically forced you to perform any other sexual act	0.013	0.046	-0.032	(0.809)	0.029	-0.075	0.104	(0.327)
	(0.005)	(0.137)			(0.006)	(0.109)		
Sexual violence: Forced you with threats/in another way to perform sexual acts	0.013	-0.158	0.172	(0.155)	0.019	0.023	-0.004	(0.970)
	(0.005)	(0.123)			(0.005)	(0.098)		

Notes: The direct report column reports the prevalence estimate derived from the direct report questions of violence in the last 13 months; the list experiment column reports the prevalence estimate derived from an analysis of the list experiment question, in which the respondent's answer is regressed on a binary variable for treatment list, conditional on woreda fixed effects and with standard errors clustered at the kebele level. The final column reports the p-value for a test of equality across the direct report and list experiment estimates.

Table A7: Subsample of respondents surveyed by male and female enumerators

	Male enumerator sample				Female enumerator sample			
	Direct report	**List report**	**Difference**	**p-value**	**Direct report**	**List report**	**Difference**	**p-value**
Emotional violence: Said something to humiliate you	0.035	-0.098	0.132	(0.210)	0.047	0.158	-0.111	(0.328)
	(0.007)	(0.106)			(0.008)	(0.121)		
Emotional violence: Threatened to hurt or harm you	0.018	-0.043	0.061	(0.570)	0.022	-0.079	0.101	(0.375)
	(0.005)	(0.110)			(0.006)	(0.115)		
Emotional violence: Insulted you or made you feel bad about yourself	0.083	0.066	0.017	(0.865)	0.124	0.123	0.001	(0.994)
	(0.010)	(0.104)			(0.014)	(0.122)		
Physical violence: Pushed you, shook you, or thrown something	0.04	0.015	0.025	(0.823)	0.035	0.16	-0.124	(0.321)
	(0.007)	(0.112)			(0.007)	(0.128)		
Physical violence: Slapped you	0.044	0.098	-0.053	(0.642)	0.064	-0.051	0.114	(0.320)
	(0.008)	(0.117)			(0.010)	(0.117)		
Physical violence: Twisted your arm or pulled your hair	0.024	-0.05	0.074	(0.451)	0.028	0.019	0.009	(0.933)
	(0.006)	(0.099)			(0.007)	(0.110)		
Sexual violence: Physically forced you to have sexual intercourse	0.019	-0.141	0.16	(0.118)	0.049	0.253	-0.205	(0.104)
	(0.005)	(0.105)			(0.008)	(0.131)		
Sexual violence: Physically forced you to perform any other sexual act	0.014	0.038	-0.024	(0.845)	0.031	-0.043	0.074	(0.515)
	(0.004)	(0.125)			(0.006)	(0.116)		
Sexual violence: Forced you with threats/in another way to perform sexual acts	0.007	-0.084	0.091	(0.331)	0.027	0.027	0	(0.997)
	(0.003)	(0.096)			(0.007)	(0.121)		

Notes: Panel A reports analysis for respondents surveyed by a male enumerator; Panel B reports analysis for respondents surveyed by a female enumerator. The direct report column reports the prevalence estimate derived from the direct report questions of violence in the last 13 months; the list experiment column reports the prevalence estimate derived from an analysis of the list experiment question, in which the respondent's answer is regressed on a binary variable for treatment list, conditional on woreda fixed effects and with standard errors clustered at the kebele level. The final column reports the p-value for a test of equality across the direct report and list experiment estimates.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available **here**

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
www.ifpri.org

IFPRI HEADQUARTERS
1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org