🔓 OPEN ACCESS

Religious leaders can motivate men to cede power and reduce intimate partner violence: Experimental evidence from Uganda

Christopher Boyer[a,1] ⓘ, Elizabeth Levy Paluck[b,c], Jeannie Annan[d], Tvisha Nevatia[e] ⓘ, Jasper Cooper[f], Jackline Namubiru[e] ⓘ, Lori Heise[g,h] ⓘ, and Rachel Lehrer[d]

Violence committed by men against women in intimate relationships is a pervasive problem around the world. Patriarchal norms that place men as the head of household are often to blame. Previous research suggests that trusted authorities can shift perceptions of norms and create behavior change. In many settings, a compelling authority on behavior in relationships is religious leaders, who are influential sources of information about proper conduct in relationships and gatekeepers of marriage, but may also uphold traditional gender roles. One way leaders exert their influence is through premarital or couples counseling courses. In this study, we test whether, if given an opportunity to offer a more progressive religious interpretation of gender roles during these courses, religious leaders could motivate men to share power and thereby reduce violence. Building on existing faith networks of Christian religious leaders in western Uganda, we conducted a large pair-matched, randomized controlled trial among 1,680 heterosexual couples in which participants were randomized to attend a 12-session group counseling course or wait-listed. We find that the program shifted power from men to women and reduced intimate partner violence by five percentage points, comparable with more intensive secular programs. These improvements were largest among couples counseled by religious leaders who held the most progressive views at baseline and who critically engaged with the material. Our findings suggest that religious leaders can be effective agents of change for reducing violence.

intimate partner violence | religion | authorities | counseling | norms

Significance

In most societies, religious leaders play an influential role in the construction of gender norms. One setting in which they exert their influence is the premarital or couples counseling course. We hypothesized that if leaders offered a more progressive religious interpretation of gender roles during these courses, emphasizing the need for men to improve their relationships by balancing power and decision-making with their spouses, violence would reduce. We find that when Christian leaders in Uganda offered these types of courses, power shifted from men to women, and intimate partner violence decreased by five percentage points a year later. Given the ubiquity of premarital counseling within churches, the intervention has the potential to reach a massive audience.

What causes men to abuse women in romantic relationships? While the proximate trigger may vary, researchers agree that male violence is a physical manifestation of historic imbalances in power between men and women that place women in a subordinate role (1–6). Globally, this form of violence is pervasive, afflicting nearly a third of women across their lifetime, and often spikes during times of crisis, such as the recent COVID-19 pandemic (7, 8). Violence persists, in part, due to strong social norms that position men as the head of the household who are entitled to sex and to their partners' deference and money and due to norms that permit violence as an appropriate means of enforcing men's position and entitlements (9). Thus, to reduce violence, many programs aim to shift people's perception of the roles and expectations that are socially or morally normative in heterosexual relationships. To change perceptions of norms, interventions use the legal system to criminalize intimate partner violence (10) or use mass media campaigns (11–14) and community mobilization (15–19).

Research suggests that norm perception is strongly influenced by signals from influential community members—and that targeting those people is an effective channel for behavior change, even when individual attitudes or beliefs about the topic are highly stable (20–23). In many contexts, the social and moral authority of religious leaders lends substantial weight to their messages about topics, like violence and health (13, 24–27). Yet, to date, there exists little to no empirical evidence about the influence of religious authorities on changing norms regarding intimate partner violence.

In this study, we tested whether a program delivered to couples by religious leaders from within the church would motivate men to cede power and reduce intimate partner violence within heterosexual couples in rural Uganda. At the 12-mo end line in a pair-matched, randomized, controlled trial, we found that the program significantly reduced the proportion of women who reported any form of violence by their partner and significantly increased women's decision-making authority within the household.

There are theoretical and historical reasons to expect that the church could meaningfully shape heterosexual relationship norms and behavioral patterns. As a moral and social authority on appropriate conduct, churches influence which behaviors are socially acceptable, especially in contexts where religious identification is strong. Through sermons,

edicts, and interpretation of scripture, they define prescriptive norms about behavior within sexual relationships and the day-to-day roles of men and women. While at present, church doctrine in many places describes norms that uphold traditional patriarchal structures, such norms are not static. They have been contested and changed across time and place (28).

The church's influence over community norms is in part exerted through its representatives: the pastors, priests, or other leaders who communicate norms through sermon and scripture as well as through community outreach programs, like premarital and marital counseling. For example, in many circumstances religious leaders serve as official mediators of couples' conflicts, during which leaders can either challenge or reinforce the church's traditional gender norms. In this way, the religious leaders of individual churches can function as "norm entrepreneurs." They are endowed with some of the authority of the institution but are also free to act in their own right to shape community norms. Historically, it is precisely this sort of religious leader who has helped to drive both progressive and reactionary social movements, from the civil rights and antiapartheid movements to the abolition and temperance movements (29–33).

While some religious leaders may choose to be agents of progressive normative change, inviting them as collaborators on violence prevention also poses several significant risks. First, many have provided theological justification for patriarchal norms in the past or may hold personal beliefs that run contrary to more equitable messages. They may even themselves be part of violent heterosexual relationships or model violent behaviors. This may make them less motivated to act as a norm entrepreneur or less credible if others perceive their normative messages to be inauthentic. Second, some leaders prioritize institutional practices or goals that conflict with violence prevention and equitable gender norms. For instance, religious leaders may prioritize the stability of heterosexual marriages above individual concerns of spouses, and therefore, they may discourage women from leaving abusive partners or emphasize norms that prompt harmony through female obedience.

The intervention tested in this study, Becoming One (B1), was designed to build on the potential of religious leaders as norm entrepreneurs while mitigating some of the risks. Drawing insights from a dynamic human-centered design process, B1 works within existing church programming models and aligns with the priority placed on heterosexual relationships and marriage by leaders and their congregants. Specifically, B1 is a curriculum embedded within faith-based marital counseling classes, an institutional experience sought by many heterosexual couples. Religious leaders like to offer these courses as a means to promote healthy Christian relationships and to recruit and retain congregants; they also raise funds for the church (many churches require counseling if couples want a church wedding).

We provide a motivational account of how such an intervention might reduce violence. First, we draw on the centrality of the Christian church in the lives of our participants. Couples seek religious counseling under the pressure of preexisting social norms and are motivated to engage with the materials due to the religious framing and the faith leader's authority in the community. Second, the benefits of improved relationship quality: the mutual understanding and intimacy inspired by the program increase the couple's valuation of the relationship. This improvement motivates the couple, and notably, the man to continue engaging in the counseling program's recommended practices around power sharing and dispute resolution. These practices bring about more equal power dynamics, such as joint rather than male-dominated

financial planning. The decrease in power inequality leads to a reduction in violence.

The B1 curriculum does not address violence directly, focusing instead on improving couples' relationships in ways that benefit both women and men. The content of the curriculum uses the language, symbols, and authority of the church to identify its recommendations for new behaviors and norms as a Christian. For example, it 1) provides alternative interpretations of scriptural passages often thought to justify male dominance, 2) creates a new aspirational identity for heterosexual Christian couples that is based in equality and trains them in the requisite skills to achieve greater power sharing, and 3) provides opportunities for the religious leader to model and socially reinforce new behaviors and norms in front of a group of Christian couples (classes are held, as is traditional for marital counseling in many contexts, for groups of several couples at once).

Our findings support the view that intimate partner violence can be reduced by those with moral authority in their communities. Entrenched patterns of behavior based on seemingly "natural" categories of identity, like man and woman or husband and wife, are changeable when individuals receive messages from these moral authorities couched in the same symbology, language, and moral framework as traditional messages about gender and power. In addition, our data are consistent with the story that violence within a relationship can be changed without directly addressing it but by addressing behaviors related to violence and improving the quality of the relationship. By teaching good communication, shared decision-making, and emotional regulation, it is possible to mitigate power imbalances and conflict.

In addition to providing an empirical answer to the question of whether religious leaders can be agents of normative and behavioral change in heterosexual relationships, we also ask what kind of religious leader can be an agent of change. Unlike previous studies, which typically randomize violence prevention programming at the community level, our blocked wait-listed design permits the estimation of program effects for each religious leader. We draw on a rich set of data on preintervention characteristics of the leaders (e.g., ideology) and on their implementation of the program (e.g., from random audits of the sessions) to explore important sources of heterogeneity in this approach. We find the positive effects on congregants to be concentrated among those in classes with religious leaders who held more progressive views on gender roles and the acceptability of violence, and who more closely followed the norms of the curriculum.

Materials and Methods

Intervention and Experimental Design. This study measures the effect of the B1 program, designed by the International Rescue Committee's (IRC) Airbel Impact Laboratory, through a pair-matched, randomized controlled trial with 3,360 men and women in monogamous heterosexual relationships and with 140 religious leaders (mainly catechists, pastors, and priests) (*SI Appendix*, Table S1 has baseline characteristics) identified by the implementing partner, World Vision, in three districts[*] in western Uganda. In mid-2018, those religious leaders identified 2,561 couples who were eligible and interested in the program and then, together with trained enumerators, invited them to participate in the program. In September 2018, in partnership with Innovations for Poverty Action (IPA), we conducted a baseline survey with 1,680 couples randomly sampled from across the 140 congregations. During the survey, both men and women were separately invited to participate in the research and provided written informed

---

[*]Specifically, these districts are Kamwenge (population = 414,454), Kagadi (population = 430,200), and Kakumiro (population = 473,400); population figures are all projections for 2019 (*SI Appendix*, Fig. S1 shows a map of the study communities).

consent. We formed those couples into 840 pairs by matching them within their congregation on baseline levels of self-reported violence in their relationship. Within each pair, we randomized one couple to begin the 12-session program immediately (October 2018) and the other to begin it in December 2019 (see flow diagram, *SI Appendix*, Fig. S2). Further methodological details, including the end line questionnaire, the B1 curriculum, and the preanalysis plan, are included in *SI Appendix*.

B1's primary aims are to prevent intimate partner violence and to shift power toward women. It attempts to do so by teaching relationship skills: communication, emotional regulation, shared control over financial resources, financial planning, and sexual consent and pleasure. Religious leaders are trained for 2 d and given an instructional guide for themselves and their couples as well as smartphones with video lessons to prepare for each session. They were also signed up for a WhatsApp group with other participating religious leaders in their district. Sessions are intended to be delivered by religious leaders to groups of four to seven couples for 12 weekly 90-min sessions. The sessions are organized thematically and are meant to be participatory, with the couples often asked to complete interactive activities together. Each couple is also given a guidebook with additional home practice activities that they are asked to complete and report back on during subsequent sessions. The guide book and other materials showcase aspirational couple identities through vivid illustrations of Ugandan couples. Lessons are justified and reinforced through Christian teachings, including reinterpretations of biblical passages commonly cited as justifications for women's subservience to men. For example, the well-known teaching "It is a woman's duty to obey and submit to the man" is challenged with Ephesians 5:21, "Submit to one another out of reverence for Christ," along with an explanation of the Bible's validation of the submission of both husbands and wives. The curriculum does not explicitly address intimate partner violence, which among other things, helps to mitigate concerns about socially desirable responses.

We measured both the treatment and control couples' outcomes in two follow-up surveys conducted in April 2019 and November 2019, roughly 6 and 12 mo after the counseling sessions had started for the treatment couples. In addition to the quantitative measurements, we also conducted qualitative focus groups and individual interviews. These were held after the first follow-up survey both with religious leaders ($n = 23$) and with couples who had completed the program ($n = 12$). The pair-matched, couple-randomized design optimizes power; the SEs in our four main analyses suggest we are able to detect a minimum effect of between 0.03 and 0.12 of a control SD at 80% power[†] (34).

Ethics. Due to the sensitivity of conducting research on violence committed by an intimate partner, we took additional measures to minimize participation risks to staff and respondents. Constraints do not allow for a full discussion here, so we direct readers to *SI Appendix*, in which we describe in detail the steps taken to protect research participants' rights as well as the ethical considerations involved in the measurement and randomization of our study. Briefly, we did not anticipate that any harm would come to women due to their participation in the trial. We maintained regular contact with participants, and a referral system was set up for participants in distress. Information about local support services was provided to all participants regardless of whether they disclosed experiencing violence, and participants were gender matched with interviewers. All project materials, including the research protocol, survey instruments, and consent forms for the baseline and end line, were reviewed and approved by the institutional review board (IRB) at IPA (protocol no. 14916), by the Mildmay Uganda Research and Ethics Committee (MUREC, protocol no. REC REF 0508 2018), and by the Ugandan National Council for Science and Technology (protocol no. SS4782).

Estimation and Inference. All main analyses were conducted in line with the preanalysis plan, which specified not only which estimators and tests we would employ, but also how we would code outcome variables and transparently record the data cleaning process. We verified experimental balance at the baseline across all measured covariates through both joint and individual

tests of equality of means (*SI Appendix*, Table S2). Response rates at midline and end line were both high, with 97.8% of couples surveyed. We found no evidence that attrition was related to treatment assignment (P value = 0.62) or that the composition of the sample differs by treatment status due to attrition (P value = 0.34), and results were robust to trimming-bound, extreme value–bound, and doubly robust inverse probability of attrition weighting approaches (*SI Appendix*, Figs. S3 and S4 and Tables S3 and S4). Compliance with treatment assignment was verified using random audits, attendance logs, and photos. Noncompliance was less than 17% in treatment group and 1% in control group. Effects among those who complied with assignment were first nonparametrically bounded and then estimated using both instrumental variable– and principal stratification–based approaches in *SI Appendix*, Tables S5–S7.

We report intent-to-treat (ITT) effects for all analyses in the paper. In our primary analysis, our estimator is a least squares regression that conditions on an indicator for the treatment assignment, fixed effects for the pair blocks, and covariates selected through a cross-validated lasso regression (*SI Appendix*). To maximize efficiency, covariates are mean centered, and they are interacted using the approach described in ref. 35 and implemented in *estimatr* (36). The regression model we estimate at the couple or individual level can be written as

$$Y_{ij} = \gamma_j + \tau Z_i + \beta^T(\mathbf{x}_i - \bar{\mathbf{x}}_i) + \delta^T Z_i(\mathbf{x}_i - \bar{\mathbf{x}}_i) + \varepsilon_i,$$

where γ_j is a fixed effect in the jth matched pair block, τ is the ITT effect evaluated at the multivariate mean of the covariates, Z_i is an indicator of assignment to the program, $\mathbf{x}_i - \bar{\mathbf{x}}_i$ is a vector of mean-centered covariates, and ε_i an error term. The index i indicates individuals or couples in individual- or couple-level analyses, respectively. We impute item-level missing values.[‡] In all couple-level analyses, we calculate standard errors (SEs) using a heteroskedasticity-robust (HC2) estimator (38). In all individual-level analyses, we calculate cluster-robust (CR2) SEs (39). Decisions about the significance of effect sizes rely on nonparametric P values calculated using randomization inference (40–42). For our four main analyses, we account for multiple comparisons using a randomization-based procedure described in *SI Appendix*.

Results

Our preanalysis plan specified analyses for four primary outcomes that would constitute the principal confirmatory findings of the study: violence, the balance of power and decision-making, communication, and sexual consent and autonomy. To maximize power and minimize the number of hypothesis tests conducted, we construct these outcomes as indices of conceptually related items and use the indices as the basis for determining statistical significance rather than testing items individually. We provided full details on the construction of indices as well as their constitutive items in *SI Appendix*.

Religious Counsel Reduced Violence and Shifted Power to Women. Table 1 reports the results of two posttreatment follow-up surveys conducted 6 and 12 mo after the start of the program. We find a reduction in partner violence, accompanied by an increase in the equality of power relations between partners and an increase in communication and conflict resolution. All effects are signed in the hypothesized direction specified in the preanalysis plan, are of moderate size relative to control means (3 to 13% changes) and SDs (0.03 to 0.20 SD effects), and are statistically significant at the preregistered $\alpha = 0.10$ level, with the exception of the results on violence at 6 mo and consent at 12 mo.

[†]We calculate the minimum detectable effect sizes (MDEs) by multiplying the SE of each regression in Table 1 by 2.8 and dividing it by the control SD of that outcome. Note that this is a conservative estimate; our P values are based on better-powered preregistered one-tailed Fisher randomization tests.

[‡]The percentage of observations with missing records is given for each of the primary outcomes in Table 1 via multivariate imputation by chained equations using *mice* (37) based on all available baseline covariates for both partners in a couple (*SI Appendix*). In general, the number of observations with imputed data varied between 2 and 6% of the sample. In *SI Appendix*, Table S8, we show that the results are similar if we consider only complete cases.

Table 1. Main results

	Any violence [0,1]		Control and decision-making [0,1]		Sexual consent and autonomy [0,1]		Communication and conflict resolution [0,1]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Wave 1: 6 mo after the start of the program								
Religious counseling	−0.026	−0.013	0.027***	0.027***	0.018**	0.012*	0.022***	0.019***
	(0.023)	(0.022)	(0.007)	(0.006)	(0.008)	(0.008)	(0.007)	(0.006)
	[0.140]	[0.268]	[0.000]‡	[0.000]‡	[0.018]†	[0.055]	[0.000]‡	[0.001]‡
Wave 2: 12 mo after the start of the program								
Religious counseling	−0.056**	−0.050***	0.017***	0.019***	0.013*	0.008	0.018***	0.014***
	(0.024)	(0.023)	(0.007)	(0.006)	(0.009)	(0.008)	(0.007)	(0.006)
	[0.014]‡	[0.009]‡	[0.002]‡	[0.004]‡	[0.058]	[0.124]	[0.007]‡	[0.001]‡
Hypothesis	—	—	+	+	+	+	+	+
Preregistered	✓	✓	✓	✓	✓	✓	✓	✓
Covariates	✗	✓	✗	✓	✗	✓	✗	✓
Block FE	✓	✓	✓	✓	✓	✓	✓	✓
Control mean (SD), 6 mo	0.38 (0.49)	0.38 (0.49)	0.65 (0.14)	0.65 (0.14)	0.61 (0.18)	0.61 (0.18)	0.63 (0.15)	0.63 (0.15)
Control mean (SD), 12 mo	0.42 (0.49)	0.42 (0.49)	0.62 (0.14)	0.62 (0.14)	0.62 (0.18)	0.62 (0.18)	0.63 (0.15)	0.63 (0.15)
Observations	1,680	1,680	1,680	1,680	1,680	1,680	1,680	1,680
Imputed, %	2.6	2.6	5.4	5.4	5.8	5.8	4.4	4.4

Estimates of the ITT effects of the B1 couples counseling program on preregistered primary outcomes at first and second survey waves are shown. The dependent variable in columns 1 and 2 is an indicator that takes the value of one if the woman reported experiencing any form of physical or sexual violence since the previous wave and zero otherwise. Columns 3 to 8 are composite indices of control and decision-making questions (columns 3 and 4), sexual consent and autonomy questions (columns 5 and 6), and communication and conflict resolution questions (columns 7 and 8), respectively. These indices are formed by scaling each item to be between zero and one and then, taking the arithmetic mean. Inference about program effects is based on nonparametric randomization inference P values using 10,000 simulated draws from the randomization distribution. Hypotheses for all primary outcomes are one tailed with the prespecified direction shown. Adjusting for multiple comparisons using the procedure described in *Materials and Methods* suggests that test-wise α levels of 0.017 and 0.03 are necessary to achieve family wise error rates (FWERs) of 0.05 and 0.10, respectively; effects that meet these FWER thresholds are denoted with crosses. Regression specifications in columns 2, 4, 6, and 8 include baseline covariates and treatment by covariate interactions as per the estimator described in Lin (35). All specifications include pair-blocked fixed effects to account for the randomization procedure used. Heteroscedasticity-consistent robust standard errors (HC2) for all specifications are shown in parentheses. The simulated randomization distribution under the null for each outcome is shown in *SI Appendix*, Fig. S5. *$P < 0.1$; **$P < 0.05$; ***$P < 0.001$; †significant at FWER = 0.10; ‡Significant at FWER = 0.05, FE, fixed effect.

In order to obtain a family-wise error rate of 5 or 10%, our simulations suggest we should apply a test-wise α (false-positive threshold) of 0.017 or 0.030. By that standard, the effects on violence (12 mo), the control and decision-making index (6 and 12 mo), and the communication and conflict resolution index (6 and 12 mo) remain statistically significant after accounting for multiple comparisons, while the effects on the consent index at both time points are not robust to this correction.

Turning first to the results on violence, we find no statistically significant evidence that B1 reduced the probability that a man committed violence against his partner 6 mo after the program started ($\hat{\tau} = -0.013$, $P = 0.268$), but it did significantly reduce the probability of violence at 12 mo ($\hat{\tau} = -0.050$, $P = 0.009$). We believe that this pattern is consistent with the idea that changes were not immediate but developed over the course of the program as couples engaged with the content. Taken at face value, the estimated five percentage point reduction at the second follow-up implies that among the 840 couples randomized to treatment, the program prevented male violence against women in ~42 couples. Fig. 1, *Upper Left* suggests that the reductions in violence are driven by decreases in the proportion of women reporting their partner coerced sex or raped, punched, or pushed them (the forms of violence most reported at baseline). In secondary analyses, we find that for women in relationships where violence persisted, modest improvements may have occurred; the program reduced the frequency of emotional, physical, and sexual violence as well as the severity and the proportion of possible acts experienced by women (*SI Appendix*, Tables S9–S11). We also see reductions in the number of women reporting hitting their partner ($\hat{\tau} = -0.010$, $P = 0.020$) and reductions in reported disciplinary violence against children ($\hat{\tau} = -0.035$, $P = 0.053$), suggesting positive spillover effects of reduced conflict for all household members (*SI Appendix*, Tables S12 and S13).

The results presented in Table 1 are broadly consistent with the program's goal to shift power dynamics within the couple; at 12 mo, we see improvements in women's control and decision-making ($\hat{\tau} = 0.017$, $P = 0.002$) as well as nonsignificant§ changes in sexual consent and autonomy ($\hat{\tau} = 0.008$, $P = 0.124$). In Fig. 1, *Lower Left*, we show that the apparent increase in women's control and decision-making is driven principally by increased involvement of women in decision-making around their partner's finances. Consistent with this idea, we also find that couples who have been assigned to the program are statistically significantly more likely to report, in separate interviews, that they have engaged in joint financial planning ($\hat{\tau} = 2.4$, $P = 0.040$) and that they engage in less income hiding (proportion hiding income: $\hat{\tau} = -0.036$, $P = 0.020$; amount hidden: $\hat{\tau} = -4,100$ Ugandan Shillings (UGX), $P = 0.006$ (*SI Appendix*, Tables S14 and S15).

These changes represent "zero-sum" shifts in power—in the sense that women gain power at the expense of their male partners who cede power—as shown in Fig. 2, which plots heterogeneity in the effects of the program on perceptions of control by gender. In general terms, women in the treatment group appear to have experienced their participation in the program as a gain in both control and decision-making power, while men report experiencing a loss in control and decision-making power. For the question on control, the effect is $0.187 - (-0.070) = 0.257$ scale points more positive for women. For the question on decision-making, the effect is $0.132 - (-0.153) = 0.285$ scale points more positive

---

§However, as we show in Fig. 1, *Upper Right*, the nonsignificant changes in sexual consent and autonomy are primarily driven by a negative effect on men's reports of using noncoercive strategies. Indeed, in an exploratory secondary analysis in which we drop this, the only male-reported item from the index, improvements are significant and more consistent with those from the control and decision-making index ($\hat{\tau} = 0.018$, $P = 0.003$).
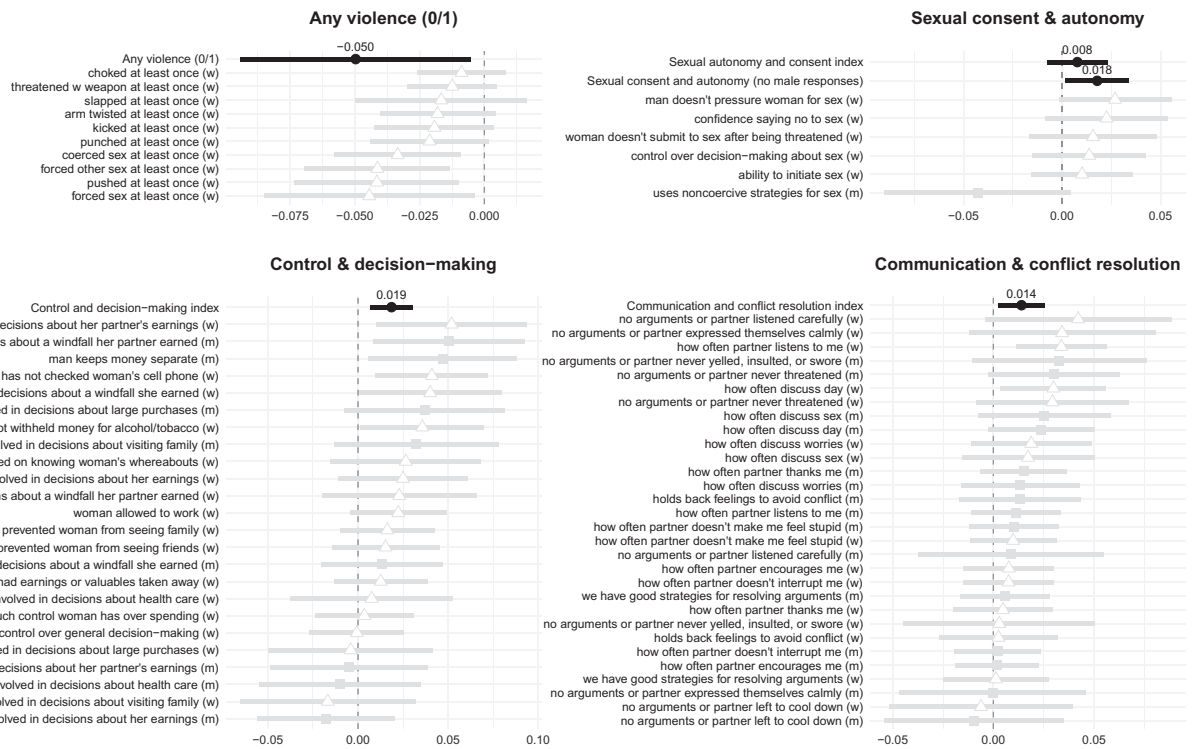
Fig. 1. Religious counseling reduces violence, increases power sharing, and improves communication. Black points represent estimated effects for preregistered main and secondary indices; gray points represent index components. Items reported by men are denoted "(m)" and are shown with squares; items reported by women are denoted "(w)" and are shown with triangles. Violence index is a binary indicator for the presence of any constitutive elements. Other indices are formed by taking the simple arithmetic mean of all items coded between zero and one. As specified in the preanalysis plan, we show constitutive items graphically but base our inference about statistical significance of the program effects on indices only. Asymptotic 95% CIs are provided for effects on all items based on HC2 robust SEs.

for women. Both are significant at the $\alpha = 0.10$ level ($P = 0.091$ and $P = 0.054$, respectively).

Both Partners Benefited from Improved Relationship. Why would men knowingly cede control to their female partners? We believe this shift is related to the increased value partners

Fig. 2. Shifts in power are zero sum (men cede power, and women gain). Preregistered heterogeneous effects of religious couples counseling on control and decision-making outcomes by gender are shown. Questions were asked on a scale from 0 to 10. Women's and men's responses are represented by triangles and squares, respectively.

place on their relationship as a result of improvements in couple dynamics that benefit both partners (i.e., "positive sum" dynamics). At a basic level, we find that B1 couples report a greater degree of trust and intimacy than those in the control group and that this is driven by a convergence in men's and women's assessments of the quality of their relationship. In secondary analyses, we find that B1 reduced reported within-couple differences in emotional closeness as well as differences in reported trust (*SI Appendix*, Tables S16 and S17). In both cases, these changes are the result of women's assessments moving closer to men's assessments of closeness and trust. We also see shared reductions in depression ($\hat{\tau} = -0.008$, $P = 0.023$), improvements in sexual intimacy ($\hat{\tau} = 0.01$, $P = 0.06$), and couples reporting spending more time together ($\hat{\tau} = 0.014$, $P = 0.001$) (*SI Appendix*, Table S18).

We believe that this increase in intimacy is a result of positive changes in communication and conflict resolution skills, as reflected in the index in Table 1. As we show in Fig. 1, the improvement in the communication index is driven primarily by the fact that women and men are more likely to report that their partner listens to them and does not interrupt, thanks them, and discusses worries with them. We also see that treated couples were more likely than their control counterparts to report that either they did not have any arguments or that they had arguments but used positive strategies for dealing with them, such as calmly expressing one's feelings, leaving to cool down, or refraining from yelling and swearing (*SI Appendix*, Table S19). We see evidence of a reduction in the frequency of arguments reported by members of the couple ($\hat{\tau} = -0.012$, $P = 0.03$) (*SI Appendix*, Table S13), but we do not see that the proportion of people who ever had arguments is decreased by the program. This suggests an improvement in conflict resolution strategies.
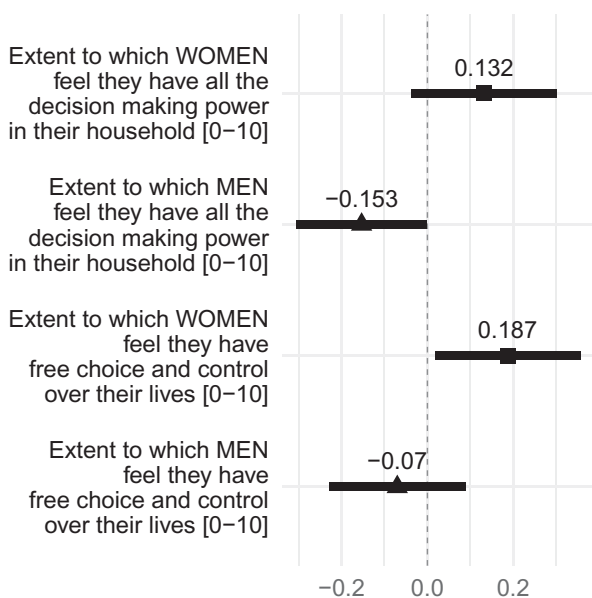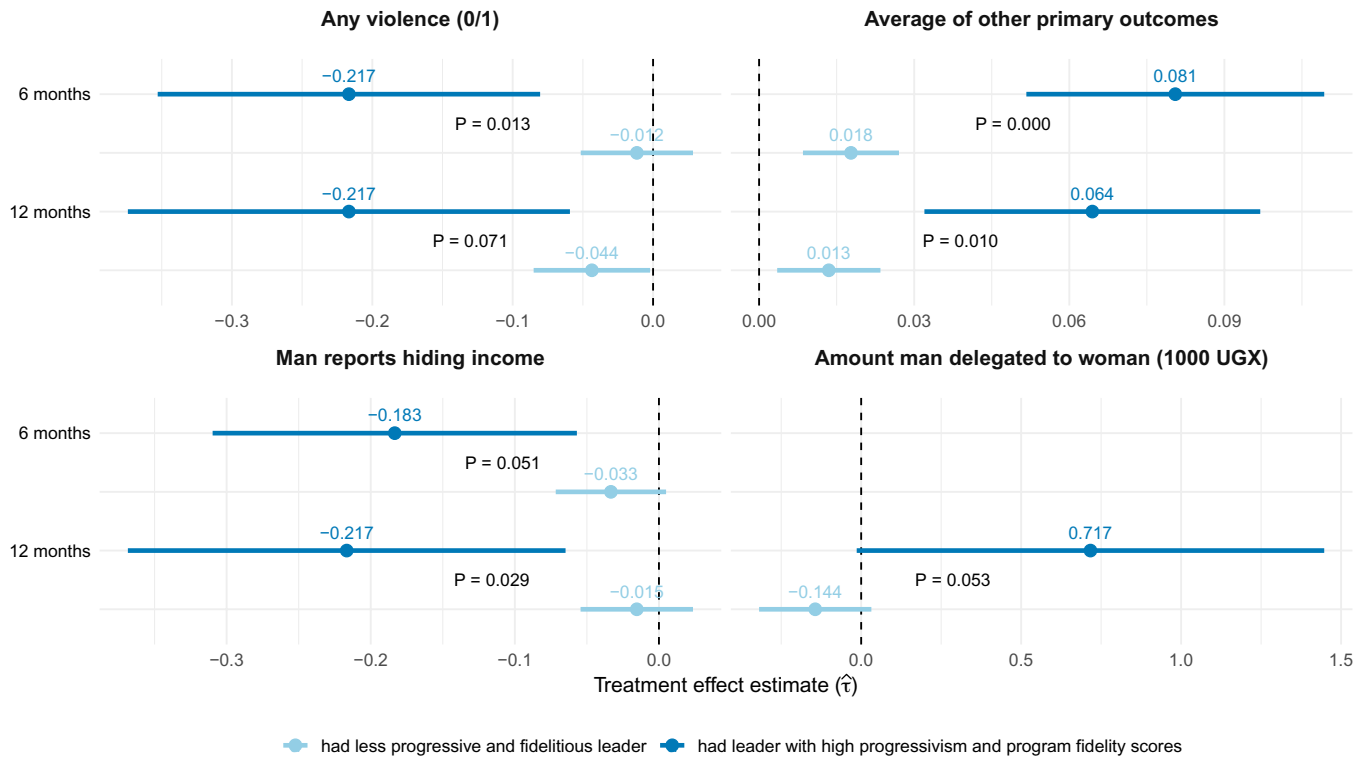
Fig. 3. Religious leaders who scored the highest in progressivism and engaged most with the program produced the biggest changes. Effects estimated among religious leaders who were in the upper 25th percentile of both "progressivism" and "program fidelity" factors (dark blue) vs. among all other leaders (light blue) are shown. Heterogeneity is assessed via the Wald test of the product of the treatment indicator and the indicator of progressive leader (*P* values are shown).

Results Varied across Religious Leaders. Given the central role of the religious leader in the B1 program, both as a norm entrepreneur and more practically, as the program facilitator and organizer, a natural question is whether the modest but significant improvements noted above are shared broadly or if some religious leaders were more effective than others. We find evidence for heterogeneity in effects across religious leaders for all four primary outcomes ($P < 0.001$ for all).

What explains the differences in success across religious leader groups? Using a confirmatory factor model, we aggregate data from routine monitoring, random audits of sessions, and preintervention interviews to describe religious leaders along four latent axes: 1) program fidelity or how well the religious leaders stuck to the curriculum and communicated the progressive message of the program, 2) progressivism or their preintervention views on gender and gender-based violence, 3) competence or their experience and stature within the church community, and 4) attendance, specifically the rate of congregants' attendance to their sessions. We construct a summary measure of religious leader performance by summing the factor scores and binning into the lower 25th percentile, the middle 50th percentile, and the upper 25th percentile to form "low"-performance, "middle"-performance, and "high"-performance groups. *SI Appendix* has further details.

In an exploratory secondary analysis, we find evidence that a combination of two of the factors—program fidelity and progressivism—most clearly defines differences in effects achieved by religious leaders (*SI Appendix*, Figs. S6 and S7 and Table S20). Fig. 3 shows that a minority of religious leaders in the highest quartile of program fidelity and progressivism produced significantly larger changes, while most other religious leader groups experienced more modest effects. In the panels in Fig. 3, *Right*, we observe a similar pattern; averaging all other primary outcomes, estimated effects are greatest among religious leaders with the highest program fidelity and gender progressivism scores. These findings are robust to treating the scores as continuous rather than discrete (*SI Appendix*, Figs. S8 and S9).

Given that religious leaders recruited couples themselves, one might wonder whether the results above might be caused by the most progressive and enthusiastic religious leaders having a better understanding of which couples in their community were most amenable to change. While we cannot rule this out as a possibility, we find little evidence from joint tests that couples in more progressive leaders' groups differ systematically from couples in other groups on observable characteristics at baseline ($P = 0.522$) (*SI Appendix*, Table S21).

Another theory is that religious leaders in these groups may have been more motivated and prepared to act as norm entrepreneurs because their values aligned with what they were asked to present. Indeed, we find evidence that religious leaders in these groups reported being more comfortable engaging couples in discussions about some of the more controversial material on sex ($\hat{\tau} = 0.19$, $P = 0.091$) (*SI Appendix*, Table S22), and this translated to bigger improvements in consent and autonomy and lower rates of sexual violence. This finding gestures more generally to the motivation of progressive religious leaders to engage couples with the new normative regime offered by the program. Interestingly, we find that these religious leaders were also able to do this without generating a backlash among couples, as they also achieved attendance rates that were roughly 10 percentage points higher than those of their peers ($P = 0.023$) (*SI Appendix*, Table S22).

Limitations and Alternative Explanations. In the absence of another source of information on violence, disentangling experimenter demand effects from true reduction in violence is extremely difficult. Because the B1 curriculum does not mention

violence explicitly,¶ here we assume that experimenter demand effects reflect a general desire to be well regarded by surveyors or for answers to reflect well on the program or religious leader.# Thus, we estimate treatment effects on outcomes that reflect positively on the respondent but which could not have plausibly been affected by the treatment. Specifically, we ask respondents how frequently, over the course of the prior year, they loaned money, fertilizer, or other things to people in their community and donated money to international nongovernmental organizations (NGOs) or charities (i.e., like World Vision or the IRC). Using our best-powered test, including lasso-selected covariates and all 3,360 respondents, we see no statistically significant evidence for treatment effects on these outcomes (*SI Appendix*, Table S23). Despite the omission of violence in the curriculum, one might still wonder whether asking women about violence at baseline could have "tipped off" respondents to the focus of the study. To address this concern, we randomly excluded all questions about violence from 30% of the baseline sample. We see no statistically significant evidence that being asked violence questions at the baseline decreases the probability of women reporting any violence at the hands of their partners at the end line. If anything, our effect estimates are more consistent with the converse: that those who were asked questions about violence at the baseline had a slight increase in the probability of reporting violence at the end line (*SI Appendix*, Table S24).

Program participation may have changed respondents' very definition of violence—for example, by defining violence more narrowly, which would cause an apparent reduction in violence without any true underlying change in behavior. However, we see that those who went through the program come away with more expansive definitions of violence, which we measure by asking respondents if they consider a description of a husband forcing his wife to have sex when she doesn't want to a form of "violence." On average, respondents are more likely to include this behavior in their categorization of violence than their peers in the control group. This effect is small (about 1.2 percentage points) and not statistically significant at the 5% level ($P = 0.09$) (*SI Appendix*, Table S26). Still, this finding might suggest that point estimates regarding reductions in violence should not be interpreted literally.

Despite these additional survey checks, one may still wonder whether these reported changes reflect true changes in behavior. To address this limitation, at the 12-mo follow-up, we designed a laboratory-in-the-field behavioral game meant to mimic one of the reported changes at 6 mo: the shift in decision-making over men's finances to women and the reduction in men's income hiding. In a preregistered procedure, men were given a small "windfall" of 10,000 UGX with which they could buy a bundle of goods and were given the option to decide themselves how to spend the money or delegate some money to their partner. While on the whole, we do not find evidence that men delegated more to their partners (*SI Appendix*, Table S25), as shown in Fig. 3 if we compare effects among the high-performing religious leaders and their peers, men who had high-performing leaders did delegate significantly more to their partners than their peers.

Putting aside the question of whether reductions in violence constituted true changes in behavior, it is also possible that

alternative mechanisms beyond those suggested here may have been more influential. A common theory of antiviolence programs is that by addressing individual attitudes and beliefs, violence may be reduced. Therefore, it could be possible that direct attitudinal change may have been responsible for violence reductions rather than the pathway described above. While we do find evidence that attitudes may have shifted as a result of the program (*SI Appendix*, Table S26), these changes were small (about one percentage point) and at least in the case of violence attitudes, did not persist at 12 mo. Similarly, it could be that peer effects, specifically an increased willingness to intervene in cases of violence, could drive the results. However, we do not find evidence for this (*SI Appendix*, Table S27). Finally, it is possible the intervention may have encouraged religious leaders to get more directly involved in couples' relationships or increased engagement with the broader church community. However, our findings suggest that treatment couples were not more likely to seek out the counsel of the religious leader to resolve problems, nor did their attendance at religious services or church activities increase (*SI Appendix*, Table S28).

Discussion

In the present study, a curriculum embedded in a couples' marital counseling program enjoyed by church members and religious leaders alike used religious rationales to teach couples how to build less gender-traditional relationships. We find that this approach, which intentionally skirted the topic of violence, led to lowered violence, greater power sharing, and more closeness between the members of participating couples after 12 wk of coursework compared with couples waiting for their turn at counseling.

Previous work emphasizes the importance of reducing power inequality in order to reduce male violence against women, with many interventions seeking to compel men to cede power through the use of a social, legal, or economic threat. The present intervention invites men to cede power voluntarily. Consistent with prespecified expectations, for example, we see evidence of increased partner equality with respect to how financial decisions are made. This effect appears to represent a zero-sum change; men experience the shift as a loss in power, and women experience the shift as a gain in power. Some theories of male dominance predict that the experience of a loss of power can provoke violent male backlash either as some expressive response or in an instrumental attempt to regain control (45–48). However, the reduction in power inequality we observe was accompanied by a reduction in the frequency and types of violence that women report experiencing. We theorize that we do not see backlash effects and by contrast, that men felt motivated to cede power due to the benefits they enjoyed from the increase in positive relationship dynamics, achieved through the counsel of a trusted religious leader from their community. Consistent with these expectations, we find that couples enjoy better time together and report less depression. We think the lack of a backlash may also be related to the fact that the program was branded as one benefiting the couple (as opposed to being one to just benefit women). We do not find evidence that the effects are driven by experimenter demand or by a range of alternative explanations put forth in the literature, including peer effects and attitudinal change.

The estimated effects are in the middle of the range compared with other researched intimate partner violence prevention programs (49) but offer an opportunity to realize benefits at great scale given the relatively short duration at 12 sessions, minimum religious leader training required at less than 2 days, and relatively low cost at 91 US Dollars (USD) per person. Most important

---

¶For safety reasons, religious leaders were provided guidance on how to respond if approached about a specific case of violence. This included reviewing definitions of what constitutes violence; however, couples themselves were not exposed to these materials, nor from qualitative interviews do we have any evidence that leaders brought them up.

#We note that recent work on measuring intimate partner violence suggests that more indirect questioning techniques that preserve anonymity do not necessarily perform better than direct questioning (43, 44).

to the potential for scale is the popularity of marital counseling programs in Ugandan society (and in many other Christian communities, where such premarital counseling programs are value-added programming supporting congregant families and events, like weddings, are integral to the role of churches and the way they are funded). In some denominations, programs like B1 are considered mandatory if congregants wish to be married in the church, which means that scaling this program is primarily a matter of asking religious leaders to adopt it and integrating it as part of church curriculum.

We believe that several important contextual factors position religious leaders as prime agents for social change in this setting and may influence where such interventions are likely to be successful. First, religious identification in Uganda is high, especially in rural communities: more than 80% of Ugandans attend religious services on a weekly basis and 86% rate religion as being "very important" to them (50). Second, violence and the acceptability of violence are also common: half of ever-partnered Ugandan women report experiencing physical or sexual violence from an intimate partner (51) and majorities of women and men believe some forms of physical violence to be justified in certain circumstances (52). Third, many alternative secular institutions, like the state, are comparatively weak or viewed with suspicion and therefore, may not be effective vehicles for behavior change (*cf* ref. 53).

Our findings on the heterogeneity of religious leaders' efficacy are relevant to the question of scale; leaders were most successful at reducing violence and motivating men to cede some power in their relationships when they were already ideologically in support of these goals at baseline and when they adhered well to the curriculum throughout. This means that not all religious leaders are ready to realize the full degree of success of this program, and future research should consider whether trainings or other kinds of influence programs (potentially mobilizing religious leaders to talk with one another about the importance of the curriculum's goals) or screening for particular characteristics would reduce the heterogeneity of the program's effects and potentially increase program impact.

Another question for future research is whether directly addressing topics of violence would enhance the effects of the program or if the power of the program is in part contained in its indirect approach. Some of the most popular sessions of the curriculum were sessions on sexual pleasure and autonomy. Sessions about violence could potentially decrease men's engagement in the program. However, including these topics would also supply partners with the requisite vocabulary to recognize various forms of violence, like emotional or financial violence, that are often undiagnosed in their own relationship and in others'

relationships in their community. An intervention that included these topics would also have to be careful to retain the sense that the content that is "authentically" or "credibly" a part of the religious community and may increase the reliance on finding the right religious leaders as facilitators for the success of the program.

One further distinctive contribution of this intervention is that it seeks to weaken patriarchal power inequalities within institutions—heterosexual marriage and the Christian church—that have traditionally served to uphold patriarchy. Thus, this intervention joins a genre of activism and scholarship investigating the promise of achieving change within existing structures, such as the state's security system (54) and political institutions (55, 56). What is perhaps unique about the approach of this intervention is that it uses leaders within the existing structure to change the way that members of couples relate to each other and to motivate them to balance power voluntarily. The change from within strategy seems particularly important to evaluate in contexts where there exist considerable constraints on state capacity. While some denominations and NGOs have actively worked to combat domestic violence in religious settings (including, although not limited to, Tearfund, Raising Voices, Restore, and Christian Aid), we know of no other study that experimentally evaluates opportunities for antiviolence reform within the institution of the Christian church.

Author affiliations: [a]Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA 02115; [b]Department of Psychology, Princeton University, Princeton, NJ 08540; [c]School of Public and International Affairs, Princeton University, Princeton, NJ 08540; [d]Airbel Impact Lab, International Rescue Committee, New York, NY 10168; [e]Innovations for Poverty Action, Kampala, Uganda; [f]Department of Political Science, University of California San Diego, La Jolla, CA 92093; [g]Department of Population, Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205; and [h]School of Nursing, Johns Hopkins University, Baltimore, MD 21205

1. J. C. Babcock, J. Waltz, N. S. Jacobson, J. M. Gottman, Power and violence: The relation between communication patterns, power discrepancies, and domestic violence. *J. Consult. Clin. Psychol.* **61**, 40–50 (1993).
2. L. L. Heise, Violence against women: An integrated, ecological framework. *Violence Against Women* **4**, 262–290 (1998).
3. R. Jewkes, Intimate partner violence: Causes and prevention. *Lancet* **359**, 1423–1429 (2002).
4. K. A. Yllo, "Through a feminist lens: Gender, power, and violence" in *Current Controversies on Family Violence*, D. R. Loseke, R. J. Gelles, M. M. Cavanaugh, Eds. (Sage Publications, ed. 1, 1993), pp. 47–62.
5. M. P. Johnson, Patriarchal terrorism and common couple violence: Two forms of violence against women. *J. Marriage Fam.* **57**, 283–294 (1995).
6. L. M. Sagrestano, C. L. Heavey, A. Christensen, Perceived power and physical violence in marital conflict. *J. Soc. Issues* **55**, 65–79 (1999).
7. S. Bourgault, A. Peterman, M. O'Donnell, "Violence against women and children during COVID-19–one year on and 100 papers in" (Tech. Rep., Center for Global Development, Washington, DC, 2021).
8. E. Leslie, R. Wilson, Sheltering in place and domestic violence: Evidence from calls for service during COVID-19. *J. Public Econ.* **189**, 104241 (2020).
9. L. McLean, L. L. Heise, E. A. Stern, Shifting and transforming gender-inequitable beliefs, behaviours and norms in intimate partnerships: The Indashyikirwa couples programme in Rwanda. *Cult. Health Sex.* **22**, 13–30 (2020).
10. M. Htun, S. L. Weldon, The civic origins of progressive policy change: Combating violence against women in global perspective, 1975–2005. *Am. Polit. Sci. Rev.* **106**, 548–569 (2012).
11. D. P. Green, A. M. Wilke, J. Cooper, Countering violence against women by encouraging disclosure: A mass media experiment in rural Uganda. *Comp. Polit. Stud.* **53**, 2283–2320 (2020).
12. E. L. Paluck, D. P. Green, Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in Rwanda. *Am. Polit. Sci. Rev.* **103**, 622–644 (2009).
13. G. Blair et al., Trusted authorities can change minds and shift norms during conflict. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2105570118 (2021).
14. A. Banerjee, E. L. Ferrara, V. H. Orozco-Olvera, "The entertaining way to behavioral change: Fighting HIV with MTV" (Tech. Rep. 26096, National Bureau of Economic Research, Cambridge, MA, 2019).
15. T. Abramsky et al., The impact of SASA! a community mobilisation intervention, on women's experiences of intimate partner violence: Secondary findings from a cluster randomised trial in Kampala, Uganda. *J. Epidemiol. Community Health* **70**, 818–825 (2016).
16. R. Jewkes et al., Impact of stepping stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: Cluster randomised controlled trial. *BMJ* **337**, a506 (2008).
17. P. M. Pronyk et al., Effect of a structural intervention for the prevention of intimate-partner violence and HIV in rural South Africa: A cluster randomised trial. *Lancet* **368**, 1973–1983 (2006).

18. J. A. Wagman et al., Effectiveness of an integrated intimate partner violence and HIV prevention intervention in Rakai, Uganda: Analysis of an intervention in an existing cluster randomised cohort. Lancet Glob. Health 3, e23–e33 (2015).

19. E. P. Green, C. Blattman, J. Jamison, J. Annan, Women's entrepreneurship and intimate partner violence: A cluster randomized trial of microenterprise assistance and partner participation in post-conflict Uganda (SSM-D-14-01580R1). Soc. Sci. Med. 133, 177–188 (2015).

20. M. E. Tankard, E. L. Paluck, Norm perception as a vehicle for social change. Soc. Issues Policy Rev. 10, 181–211 (2016).

21. E. L. Paluck, H. Shepherd, P. M. Aronow, Changing climates of conflict: A social network experiment in 56 schools. Proc. Natl. Acad. Sci. U.S.A. 113, 566–571 (2016).

22. M. E. Tankard, E. L. Paluck, The effect of a supreme court decision regarding gay marriage on social norms and personal attitudes. Psychol. Sci. 28, 1334–1344 (2017).

23. J. E. Dannals, D. T. Miller, Social norm perception in groups with outliers. J. Exp. Psychol. Gen. 146, 1342–1359 (2017).

24. A. Abaluck et al., Impact of community masking on COVID-19: A cluster-randomized trial in Bangladesh. Science 375, eabi9069 (2022).

25. A. Luft, Religion in Vichy France: How meso-level actors contribute to authoritarian legitimation. Eur. J. Sociol. 61, 67–101 (2020).

26. C. Adida et al., Taking the cloth: How religious appeals increase compliance with COVID-19 prevention measures. SocArXiv [Preprint] (2021). https://osf.io/preprints/socarxiv/yt3e7/ (Accessed 6 December 2021).

27. K. Vyborny, Persuasion and public health: Evidence from an experiment with religious leaders during COVID-19 in Pakistan. SSRN [Preprint] (2021). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3842048 (Accessed 6 December 2021).

28. B. A. Barr, The Making of Biblical Womanhood: How the Subjugation of Women Became Gospel Truth (Baker Books, 2021).

29. J. R. McKivigan, The War against Proslavery Religion: Abolitionism and the Northern Churches, 1830–1865 (Cornell University Press, 2018).

30. D. W. Houck, D. E. Dixon, Rhetoric, Religion and the Civil Rights Movement, 1954–1965 (Baylor University Press, 2006).

31. L. L. Hunt, J. G. Hunt, Black religion as BOTH opiate and inspiration of civil rights militance: Putting Marx's data to the test. Soc. Forces 56, 1–14 (1977).

32. J. R. Gusfield, Symbolic Crusade: Status Politics and the American Temperance Movement (University of Illinois Press, 1986).

33. D. L. Chappell, A Stone of Hope: Prophetic Religion and the Death of Jim Crow (University of North Carolina Press, 2004).

34. A. Gelman, J. Carlin, Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. Perspect. Psychol. Sci. 9, 641–651 (2014).

35. W. Lin, Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. Ann. Appl. Stat. 7, 295–318 (2013).

36. G. Blair, J. Cooper, A. Coppock, M. Humphreys, L. Sonnet, estimatr: Fast Estimators for Design-Based Inference. https://declaredesign.org/r/estimatr/ and https://github.com/DeclareDesign/estimatr. Accessed 26 July 2022.

37. S. van Buuren, K. Groothuis-Oudshoorn, mice: Multivariate imputation by chained equations in R. J. Stat. Softw. 45, 1–67 (2011).

38. J. G. MacKinnon, H. White, Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. J. Econom. 29, 305–325 (1985).

39. J. E. Pustejovsky, E. Tipton, Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. J. Bus. Econ. Stat. 36, 672–683 (2018).

40. A. S. Gerber, D. P. Green, Field Experiments: Design, Analysis, and Interpretation. (W. W. Norton & Company, New York, NY, ed. 1, 2012).

41. G. W. Imbens, D. B. Rubin, Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction (Cambridge University Press, Cambridge, United Kingdom, 2015).

42. A. Young, Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. Q. J. Econ. 134, 557–598 (2019).

43. D. O. Gilligan, M. Hidrobo, J. Leight, H. Tambet, Using a List Experiment to Measure Intimate Partner Violence: Cautionary Evidence from Ethiopia (International Food Policy Research Institute, 2021), vol. 2094.

44. D. S. Park et al., "Private but misunderstood? Evidence on measuring intimate partner violence via self-interviewing in rural Liberia and Malawi" (Working Paper 29584, National Bureau of Economic Research, Cambridge, MA 2021).

45. L. M. Bates, S. R. Schuler, F. Islam, K. Islam, Socioeconomic factors and processes associated with domestic violence in rural Bangladesh. Int. Fam. Plan. Perspect. 30, 190–199 (2004).

46. G. J. Bobonis, M. González-Brenes, R. Castro, Public transfers and domestic violence: The roles of private information and spousal control. Am. Econ. J. Econ. Policy 5, 179–205 (2013).

47. L. Heise, "What works to prevent partner violence? An evidence overview" (Tech. Rep., STRIVE: London School of Hygiene and Tropical Medicine, London, United Kingdom, 2011).

48. S. Vyas, C. Watts, How does economic empowerment affect women's risk of intimate partner violence in low and middle income countries? A systematic review of published evidence. J. Int. Dev. 21, 577–602 (2009).

49. C. Bourey, W. Williams, E. E. Bernstein, R. Stephenson, Systematic review of structural interventions for intimate partner violence in low- and middle-income countries: Organizing evidence for prevention. BMC Public Health 15, 1165 (2015).

50. L. Lugo, A. Cooperman, Tolerance and Tension: Islam and Christianity in Sub-Saharan Africa (Pew Research Center, Washington, DC, 2010), vol. 147.

51. Uganda Bureau of Statistics, "Uganda Demographic and Health Survey 2016" (Tech. Rep., Uganda Bureau of Statistics, Kampala, Uganda, 2018).

52. A. C. Tsai et al., Measuring personal beliefs and perceived norms about intimate partner violence: Population-based survey experiment in rural Uganda. PLoS Med. 14, e1002303 (2017).

53. R. A. Blair, S. M. Karim, B. S. Morse, Establishing the rule of law in weak and war-torn states: Evidence from a field experiment with the Liberian National Police. Am. Polit. Sci. Rev. 113, 641–657 (2019).

54. S. Karim, K. Beardsley, Female peacekeepers and gender balancing: Token gestures or informed policymaking? Int. Interact. 39, 461–488 (2013).

55. L. Beaman, R. Chattopadhyay, E. Duflo, R. Pande, P. Topalova, Powerful women: Does exposure reduce bias? Q. J. Econ. 124, 1497–1540 (2009).

56. P. van der Windt, M. Humphreys, R. Sanchez de la Sierra, Gender quotas in development programming: Null results from a field experiment in Congo. J. Dev. Econ. 133, 326–345 (2018).

57. C. Boyer et al., Replication data for "Religious leaders can motivate men to cede power and reduce intimate partner violence: Experimental evidence from Uganda." Harvard Dataverse. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/9TEAIV. Deposited 14 July 2022.