Changing Harmful Norms through Information and Coordination: Experimental Evidence from Somalia

Pedro Ferreira[1]
Selim Gulesci[2]
Eliana La Ferrara[3]
David Smerdon[4]
Munshi Sulaiman[5]

16 September 2023

Preliminary. Please do not circulate

Abstract

We experimentally evaluate three interventions to decrease the prevalence of infibulation, the most harmful type of female genital cutting (FGC) in Somalia: (i) correcting misperceptions about support for the practice; (ii) public declarations of one's willingness to abandon the practice; and (iii) a combination of the two. We find that on average community members overestimate the community support for infibulation. Correcting this misperception reduces the probability of infibulation by 40%, two years after the intervention. Over the same time horizon, this leads to an increase in the intermediate type of FGC (*Sunna*), but it also increases the likelihood that parents report they plan not to cut their younger, uncut daughters. The public declaration treatment does not significantly affect infibulation, except in communities where participants had high priors about community support for abandoning the practice. No synergies are observed by combining the two types of interventions.

1 Harvard University; pedro_ferreira@hks.harvard.edu.
2 Trinity College Dublin, CEPR and LEAP; gulescis@tcd.ie.
3 Harvard University, CEPR, NBER and LEAP; elaferrara@hks.harvard.edu.
4 University of Queensland; d.smerdon@uq.edu.au.
5 BRAC Institute of Governance and Development; munshi.sulaiman@brac.net.

1 Introduction

Gender norms that negatively affect women's well-being persist in many societies. Female Genital Cutting (FGC) is a particularly harmful one. FGC is the custom of removing part or all of the external female genitalia for non-medical reasons and can have dramatic consequences on women' health and human capital accumulation (UNICEF, 2022). The practice is present in 29 African and Asian countries, with more than 200 million women being cut worldwide. In Somalia, FGC is nearly universal.[1]

There are a number of possible explanations for the persistence of harmful norms such as FGC. In this paper we explore two. First, FGC may persist because of a *coordination failure*. In a setting where individual choices are interdependent, multiple equilibria may exist and one reason for the persistence of bad norms can be a failure to coordinate to the 'better' equilibrium (Brock and Durlauf, 2001; Young, 2015). In such a world, even if individuals knew that others do not like FGC, no one may want to move unilaterally due to fear of social sanctions. Second, individuals may be *misinformed*. In this case, people may privately dislike FGC but think that others approve of it, and therefore they may continue following the practice – an instance of 'pluralistic ignorance' (Jaldesa et al., 2005; Gele et al., 2015). Recent evidence suggests that misplaced beliefs about others' expectations play a significant role in driving norms, and correcting such misperceptions can have real effects (Bursztyn et al., 2020).

In order to test the extent to which these two explanations can account for the persistence of FGC, we designed and implemented a field experiment in Somalia. Our experiment involves two key components: one directed at providing information (hence correcting misperceptions) and one offering a coordination device to facilitate collective action among those who privately support abandoning FGC. The latter entails a form of 'public declaration' whereby individuals can signal their position in front of other community members. To the best of our knowledge, this is the first time an intervention to correct misperceptions is implemented to address FGC and it is also the first time a study experimentally evaluates the impact of a 'public declaration' to abandon the practice.

Our experiment was implemented in collaboration with Save the Children International (SCI). We organized meetings in 141 communities, separately for men and women. Each meeting convened 29 people on average. In these meetings, participants and facilitators discussed the benefits of ending the most harmful type of FGC, infibulation (known as "Pharaonic" circumcision in Somalia) and potential pathways to abandon the

1 Based on UN data, 98% of women and girls aged 15-49 have some form of FGC (https://www.unfpa.org/data/fgm/SO).

practice.[2] Before the discussion, participants were asked to privately vote if they supported the abandonment of Pharaonic FGC, and were subsequently incentivized to guess other participants' support. We then randomly assigned communities to four equally-sized experimental arms. In the first (T1), facilitators counted the votes and revealed the aggregate level of support among the participants in the meeting. In a second arm (T2), facilitators did not announce the outcome of the vote, but allowed participants to publicly express their support for abandoning Pharaonic FGC by stepping into a circle drawn on the floor. This was a simple way to help individuals who wanted to abandon the practice know who else was in support and coordinate. In a third group of communities (T3), both treatments were implemented, i.e., the results of the poll were announced and individuals were allowed to publicly signal their support. In the remaining communities (Control) the meeting was concluded without announcing the result of the vote or allowing participants to publicly signal their position. In total, the experiment involved 8,059 individuals (3,960 men and 4,099 women) who participated in the meetings. We tracked and re-interviewed a random sample of 3,962 participants approximately 9 months and 1.5 years after the meetings were conducted. In each survey, we collected information about the FGC status of respondents' daughters aged 0-18, respondents' attitudes towards FGC, and their beliefs about other people's attitudes.

We find that on average community members overestimate other people's support for infibulation. On average, 80% of the meeting participants voted in favor of abandoning the practice, while they guessed that 74% of the participants would do so. Correcting this misperception proved effective in reducing infibulation rates. In communities where the result of the vote was announced (T1), approximately 1.5 years after the meeting girls aged 0-18 were 1.7 percentage points less likely to be Pharaonic-cut. This corresponds to a 40% decrease over the control group mean. Over the same time period, T1 led to an increase in the intermediate type of FGC ("Sunna"). It also increased the likelihood that parents reported they planned not to cut their younger, uncut daughters in the future – suggesting potential long term effects of correcting misperceived beliefs. On the other hand, the public declaration treatment (T2) did not have any significant effects on infibulation, nor other forms of FGC, on average – though interesting heterogeneity emerged, as we discuss below. The combined treatment (T3) had similar but slightly smaller effects as T1.

Examining the mechanisms, we find that in communities where the result of the vote was announced, participants' beliefs about the community support for abandoning Pharaonic circumcision had shifted at midline. Moreover, the effects of T1 are mainly

2 In what follows, we use the terms "infibulation" and "Pharaonic" interchangeably. As we explain below, FGC in Somalia takes different forms. In addition to the Pharaonic type, there is a somewhat 'milder' form called "Sunna".

driven by participants who at baseline had underestimated the true support for abandoning the practice, consistent with Bayesian updating.

For the public declaration treatment (T2), we find some evidence that it was more effective in communities where participants had higher expectations on the level of community support for abandoning Pharaonic. In those communities, more people decided to make the public declaration and we find lower rates of Pharaonic circumcision at endline. This suggests that having high second-order beliefs may be necessary for public declarations to be effective at curbing harmful norms. T2 also led to interesting changes in communication patterns within the community: at endline, respondents in T2 communities are more likely to report that they talk to at least one other household about FGC, but the *number* of households with whom they talke is lower than in control communities. By making it public which individuals were in favor of abandoning Pharaonic FGC, T2 may thus have fostered communication, but at the same time possibly reduced the diversity of opinions to which one was exposed, possibly reinforcing conservative attitudes.

Among potential mechanisms, we also investigate the extent to which our interventions may have indirectly conveyed information on the intrinsic value of infibulation, thus leading to a shift in people's willingness to adopt the practice, for any given level of community adoption. We find some evidence pointing in this direction. In particular, T1 reduced the perceived religious value and the perceived benefits of infibulation in terms of chastity, but it did not affect the expected returns on the marriage market, nor the perceived health consequences.

Finally, using a novel survey measure, we test whether individuals' degree of conformism was affected by our interventions. In a series of hypothetical scenarios, we asked respondents whether they would cut their daughters if different proportions of the community did so. We find a positive correlation between individual propensity to cut and the hypothetical community share, consistent with social pressure. However, the *slope* if this relationship is unaffected by our treatments: T1 shifts the curve down in a parallel way, suggesting that the extent to which people cared about others' opinions did not change in response to our treatment.[3]

Our paper contributes to the broad literature on social norms and development and to a recent literature looking at female genital cutting from an economic standpoint. Within the broader literature, previous work has highlighted the persistence of gender norms (Fernandez et al., 2004; Giuliano, 2007; Fernandez and Fogli, 2009; Alesina et al., 2013) and how these norms can contribute to gender inequalities (Jayachandran, 2015). Examples of harmful gender norms include child marriage (Jensen and Thornton, 2003; Ambrus and Field, 2008), bride price (Anderson, 2007; Ashraf et al., 2020), dowry (Anderson and Bidner, 2015; Bhalotra et al., 2020) and norms that limit female labor force

3 T2 and T3 did not lead to detectable shifts in the curve.

participation (Bursztyn et al., 2020). The last study has been very influential in inspiring interventions aimed at correcting misperceptions through belief elicitation and revelation[4] . We contribute in two ways. First, we assess the relevance of the pluralistic ignorance hypothesis in the context of FGC –a particularly detrimental norm whose persistence appears staggering in the face of the great health and psychological costs it imposes, and for which the existence of information gaps may not therefore be obvious. Second, as mentioned earlier, even if one managed to realign people's beliefs, this may not be enough to induce a change if coordination problems exist. Our intervention allows to assess the relative importance of misperceptions and of coordination failures, and the potential complementarity of tackling the two jointly. Even though we do not find evidence of strong complementarity between the two types of interventions, we believe this is important methodologically, and results may be different for other types of norms.

Turning to the literature on FGC, some studies have traced the origins of FGC to historical characteristics of societies (Becker, 2018; Corno et al., 2020), while a number of papers have stressed the role played by beliefs and coordination in contributing to the persistence of the practice (Mackie, 1996; Bicchieri and Marini, 2016; Efferson et al., 2015; Novak, 2020; Platteau et al., 2018). There is limited evidence on the effectiveness of policies that aim to reduce the prevalence of FGC (e.g. Diop et al., 2004; UNICEF, 2008; Camilotti, 2016; Vogt et al., 2016). Most of these studies rely on non-experimental methods. Two exceptions are Vogt et al. (2016), who evaluate the impact of edutainment programs in changing attitudes towards FGC in Sudan, and Corno and La Ferrara (2023), who estimate the effects of an information campaign and a 'norm replacement' intervention in Sierra Leone. To the best of our knowledge, our study is the first to provide experimental evidence on the effectiveness of public declarations and of correcting misperceived beliefs on FGC prevalence.[5]

The remainder of the paper is organized as follows: Section 2 describes the conceptual framework that guides our experiment, Section 3 provides background information about the context of the study and describes the experiment, Section 4 presents the results and Section 5 concludes.

4 Yu (2023) uses a similar intervention to correct misperceptions about stigma that leads to an increase in HIV testing. Another body of literature has also discussed how elections can reveal the true support to certain topics and affect individuals' beliefs and behavior (Apffelstaedt et al., 2022; Hagemeister, 2022). Recent work has also explored belief wedges in explaining

5 A number of NGOs –the most widely known being Tostan– have implemented interventions based on community sensitization followed by public declarations of the intention to abandon FGC. Past evaluations have relied on qualitative or non-experimental methods (Diop et al., 2004; UNICEF, 2008; Cislaghi et al., 2019).

2 Conceptual Framework

There is an extensive theoretical literature on social interactions with discrete choice, beginning with the seminal work of Brock and Durlauf (2001). We provide a non-technical conceptual framework that draws on this literature and motivates our interventions. A common feature of these models is that bad equilibria can persist when individuals face a trade-off between acting according to their personal preferences and following a group norm (Soetevent and Kooreman, 2007; Smerdon et al., 2020; Andreoni et al., 2021; Duffy and Lafky, 2021). In the canonical model, N players choose between two options H and L (for "high utility" and "low utility", respectively) in a coordination game. An individual's choice value is comprised of a direct utility and a social utility, which is increasing in the share of group members conforming to the same choice. For instance, an individual i's value from a choice $\omega_i \in \{L, H\}$ may be written as:

$$V_i = u_{\omega_i} + J \sum_{j \neq i}^{N-1} \left(\mathbb{1}_{\omega_i = \omega_j} - \mathbb{1}_{\omega_i \neq \omega_j} \right) + \epsilon_{\omega_i} \tag{1}$$

The middle term represents social utility, in which i will receive a positive amount $J > 0$ for every group member who shares their choice and will lose J for every group member making a different choice. The term ϵ_{ω_i} represents i.i.d. shocks to introduce heterogeneous preferences and are typically assumed to follow a Gumbel distribution for analytical tractability.

Brock and Durlauf (2001) and others show that there can be multiple stable equilibria, in which (most) individuals coordinate on one choice, so long as conformity concerns (J) are sufficiently large relative to the difference in direct utilities ($u_H - u_L$). Equilibrium selection depends on the initial conditions: if (most of) the group starts out choosing the low-utility choice L, then the group may become stuck coordinating on this "bad" equilibrium.

Two common features of these models are worth noting. First, the social interactions literature mainly focuses on binary choice because solving for all equilibria in the multinomial choice model with more than two choices becomes increasingly complicated and requires either numerical methods or restrictive assumptions to characterize all equilibria (Brock and Durlauf, 2002, 2005; Durlauf and Ioannides, 2010; Caiado et al., 2016). In recent related work, Gulesci et al. (2021) develop a theoretical model with three choices, which is relevant for the FGC decision in Somalia.[6]

Second, models typically assume that individuals have rational (consistent) expect-

6 Using a combination of theoretical and empirical analyses, Gulesci et al. (2021) show that the intermediate action may facilitate a stepping stone towards the eventual elimination of the norm or it may become a new norm in itself. Their analysis suggests that Sunna circumcision is likely to become the new norm in Somalia.

ations.[7] While this assumption greatly simplifies the equilibrium analysis, it does not allow for the group-level phenomenon of *pluralistic ignorance* to emerge. Pluralistic ignorance refers to a situation where individuals privately reject a norm but believe others accept it, leading them to conform publicly (Katz et al., 1931). This results in a disconnect between private beliefs and public actions within a group.[8] In a lab experiment, Smerdon et al. (2020) show that uncertainty over the payoffs of others can lead to outcomes that are not predicted by the standard model of social interactions but that are consistent with pluralistic ignorance. To our knowledge, Fernández-Duque (2022) is the only model that seeks to explain the (probability of) emergence of pluralistic ignorance. In this binary-choice model, uncertainty about the group's true attitudes, group size, and the initial conditions are all key ingredients for the emergence of equilibria in which individuals conform to what they incorrectly believe to be the majority attitude.

We draw on this literature to conceptualize the persistence of the practice of FGC in Somalia and to motivate our field experiment. First, FGC is often viewed as the result of a coordination failure. In a setting where individual choices are interdependent and multiple equilibria may exist, one reason for the persistence of bad norms can be a failure to coordinate to the "good" equilibrium (Brock and Durlauf, 2001; Young, 2015). Individuals may know that others do not like FGC, but no single individual may want to move unilaterally due to fear of social sanctions. Several NGOs (of which Tostan is the most widely known) have implemented interventions based on community sensitization followed by "public declarations" of abandonment of FGC. These interventions rely on the assumption that FGC is a norm that persists due to a coordination failure and if a sufficiently large number of people in the community can be convinced to commit not to cut their daughters, this can shift the community to a "no cutting" equilibrium. The declarations also reveal the identities of (would-be) "norm-breakers", which may further facilitate a transition if individuals interact locally in small-knit groups (Ellison, 1993; Kreindler and Young, 2013) or if it assists with communication of true preferences (Smerdon et al., 2020). Despite the popularity of this approach, there had been no rigorous evidence on its effectiveness up until now. Past evaluations have relied on qualitative or non-randomized evaluations with small sample sizes and literature has challenged the

7 An alternative setup involves individuals being randomly matched into pairs, with perfect, common knowledge of payoffs and preferences (e.g. Andreoni et al., 2021). The equilibrium predictions are qualitatively the same.

8 Social psychologists have documented many applications in which pluralistic ignorance is linked to the persistence of bad social norms, such as alcohol and drug use (Prentice and Miller, 1993; Barman-Adhikari et al., 2017), schoolyard bullying (Sandstrom et al., 2013), racial attitudes (O'Gorman, 1975), health care (Brener et al., 2015; Karaffa and Koch, 2016), and gender norms (De Souza and Schmader, 2022; Miyajima and Yamaguchi, 2017; Van Grootel et al., 2018; Bursztyn et al., 2020).

understanding that FGC is a coordination failure.[9]

Second, individuals may be misinformed about the support for the norm. They may privately dislike FGC but think that others approve of it, and therefore they continue doing it, leading to a state of pluralistic ignorance (Jaldesa et al., 2005; Gele et al., 2015). Misperceptions around gender norms are pervasive throughout the world (Bursztyn et al., 2023). Bursztyn et al. (2020) show that majority of young married men in Saudi Arabia support women working outside the home but they underestimate support by other similar men and correcting these misperceptions increases men's support for their wives to work outside the home. Guided by this evidence, we designed an intervention to explore if a similar approach based on correcting misplaced beliefs about support for FGC may affect attitudes towards FGC and ultimately reduce its prevalence.

3 Background and Data

3.1 Context

FGC is the practice of cutting or removing part of the external or internal female genitalia for non-medical reasons. The custom is present in 29 African and Middle Eastern countries (Camilotti, 2016), and in some of these countries FGC it is almost universal: the share of cut women is 99% in Somalia, 96% in Guinea, 93% in Djibouti and 91% in Egypt and Sierra Leone (Yoder et al., 2013; Somalia National Bureau of Statistics, 2020). An estimated 200 million women are cut worldwide (UNICEF, 2016) and every year, 3 million female infants and children are at risk of undergoing FGC (Spisma et al., 2012).

FGC is a harmful practice along several dimensions. It leads to serious health consequences, both at the time of cutting (e.g., excessive bleeding and increased mortality) and in the long run (e.g., birth-related complications and vaginal and urinary problems). It is estimated that in the cohort of 15-year-old women of six African countries[10] , a loss of 130,000 life years is expected due to FGC's association with obstetric hemorrhage (Adam et al., 2010). The most harmful form of FGC (infibulation) also poses the risk of later surgeries, like deinfibulation and reinfibulation, to allow for sexual intercourse and childbirth, further increasing both short-term and long-term risks (WHO, 2023). In Somalia, 63% of women who underwent FGC had their genitalia sewn closed through infibulation (UNICEF, 2013). Moreover, FGC is associated with lower educational achievement of girls because girls are often pulled out of school after the practice, either due to health complications or because they are seen as having "transitioned" to adulthood. Finally, given that FGC is generally performed on young girls without their consent, it is also a

9 In a nutshell, Efferson et al. (2015) argue that, were FGC a coordination problem, there would only be communities with extremely high or extremely low cutting rates. However, cutting rates in communities are not discontinuously distributed.

10 Burkina Faso, Ghana, Kenya, Nigeria, Senegal, and Sudan.

human rights violation.

FGC can take different forms. The World Health Organization classified FGC into three main types: *type I* is the partial or total removal of the clitoris and/or the prepuce (clitoridectomy); *type II* is the partial or total removal of the clitoris and the labia minora, with or without excision of the labia majora (excision); *type III*, also known as infibulation, is the narrowing of the vaginal orifice with the creation of a covering seal by cutting and appositioning the labia minora and/or the labia majora, sometimes through stitching, which may or may not include the excision of the clitoris. A fourth type includes "all other harmful procedures to the female genitalia for non-medical purposes, e.g. pricking, piercing, incising, scraping and cauterizing the genital area", but traditionally the main types considered are types I to III. In Somalia, FGC is divided into two broad categories: "Sunna" (types I–II) and "Pharaonic" (type III). Historically, Pharaonic circumcision was the dominant type in Somalia, practiced almost universally (Abdalla, 1982). In recent years, there has been a shift away from Pharaonic and a corresponding rise in Sunna circumcision (Gulesci et al., 2021).

3.2 Interventions

Motivated by the conceptual framework described in Section 2, we designed a field experiment to experimentally evaluate two interventions to mitigate the prevalence of FGC in Somalia: (1) providing information on community support to the practice and, hence, correcting misperceptions and (2) providing a coordination device in the form of public declarations that can facilitate collective action among those who support the abandonment of FGC. For the implementation of these interventions, we collaborated with Save the Children International (SCI) and its network of local contacts in Somalia.

A key challenge was proposing an intervention based on misperception and coordination in a context with wide support to FGC, like Somalia. As described in Section 3.1 above, girls in Somalia can go through one of two forms of FGC: infibulation (referred as "Pharaonic circumcision") or clitoridectomy/excision (both referred as "Sunna circumcision"). Even though both forms are present in Somalia, with 74% of adult women in our sample being Pharaonic cut and 24% Sunna cut, the two types differ in some respects. On top of being much more harmful than Sunna, Pharaonic circumcision has nearly no religious support and has been a declining practice since the mid-1990s. On the other hand, Sunna circumcision is perceived to have religious grounds and has broad support in the population. Therefore, an intervention targeting all forms of FGC (including Sunna circumcision) that aims at correcting misperceptions and promoting coordination could risk further reinforcing the harmful norm, revealing the broad support for it. Therefore, the project's intervention targeted only the most harmful form of FGC (Pharaonic circumcision) where correcting misperceptions and allowing for coordination is more likely

to have a positive impact.

We partnered with SCI to organize community meetings in 141 communities located in Somaliland and Puntland regions of Somalia (in the districts of Badhan, Burao, Erigabo, Galdagob, Galkayo and Hargeisa). Appendix Figure A.1 displays the locations of the communities included in the study. The meetings were facilitated by local personnel trained in interpersonal communication on sensitive topics to discuss the benefits of ending the most harmful type of FGC (infibulation, or Pharaonic circumcision) and potential pathways to abandon the practice. In each community, meetings were organized separately for men and women, each gender-segregated meeting convened 29 people on average. In total, the intervention reached 8,059 individuals who participated in those meetings.

During the meetings, participants were asked to state if they favored the discontinuation of infibulation using anonymous voting cards. Voting cards were truthfully anonymous and the research team cannot trace back respondents' individual votes. Before the poll, participants were incentivized individually to estimate what fraction of the group members would respond positively. Accurate guessing of the share in support was incentivized by telling participants that those whose estimates were closest to the true rate would receive a payoff in the form of mobile money.[11] To identify the specific mechanisms through which the interventions may lead to impact, we randomized the 141 communities into four groups. The randomization was stratified by district (6 strata).

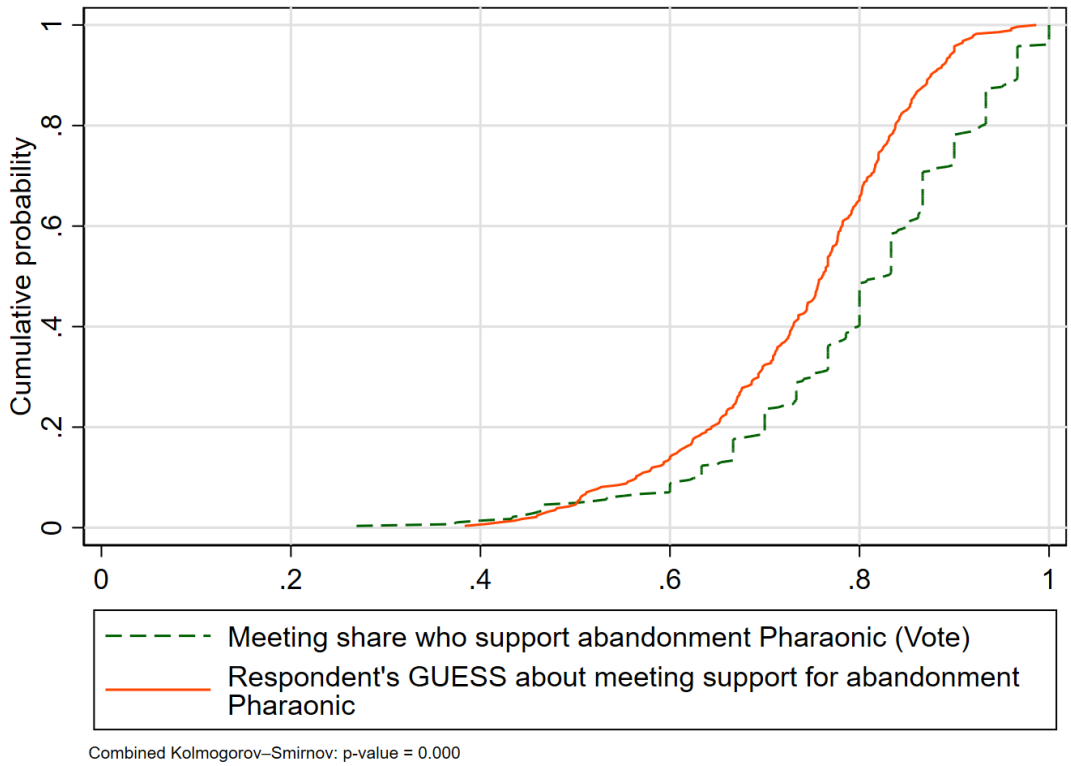
In 36 randomly selected villages, at the end of the meeting, the facilitator announced the result of the poll, thus revealing the "true" share of community members who were in favor of abandoning FGC (Treatment 1). This revelation should correct individuals' misperceptions of the attitudes in their communities, which may affect cutting decisions, thus addressing the pluralistic ignorance mechanism. In 34 randomly selected villages, at the end of the poll, we did not announce the results of the poll but rather allowed respondents to publicly express their support to abandon the practice, by stepping into a circle drawn on the floor (Treatment 2). After the poll was conducted, facilitators drew a circle on the ground and asked people who believed FGC should be abandoned to step inside the circle, while those who believed FGC should have remained outside it. The goal of this exercise was to help individuals who favored the discontinuation of FGC but did not want to move unilaterally (for fear of social sanctions) to coordinate to their preferred equilibrium more easily, by lowering the cost of communication and coordination. Thus, it addressed the coordination failure mechanism. In other 36 randomly selected villages, we both announced the result of the poll and allowed respondents to publicly express their support by stepping into the circle draw on the floor (Treatment 3). In this case, facilitators first announced the results of the poll and then conducted the coordination

11 The amount of the payoff was equivalent to 20 USD.

exercise of public declarations by stepping into the circle drawn on the floor.[12] The remaining 35 villages were left out as a control group, which we will use to identify the effects of the treatments. In control communities, the meeting and the vote were conducted similar to Treatment communities but the meeting was concluded without announcing the result of the vote or running the coordination exercise. Appendix Table A.1 summarizes the design of the experiment.

Data collected during the meetings shows evidence in line with pluralistic ignorance. On average, respondents guessed that 74% of other meeting participants would support abandoning Pharaonic circumcision, whereas the average level of support across all meetings was 80%. Figure 1 displays the cumulative distribution functions (CDF) for the share of respondents who support the abandonment of Pharaonic circumcision (i.e., the result of the vote, represented by the dashed green line) and the respondents' average guess per meeting about the share who supports it (the red line). The CDF for the share of respondents who support the abandonment of Pharaonic circumcision lies to the right of the CDF for respondents' guess, suggesting that respondents underestimated the support for abandonment of Pharaonic circumcision.[13]

Figure 1: Beliefs about community support vs. Actual support for abandoning Pharaonic FGC: Evidence on pluralistic ignorance at baseline

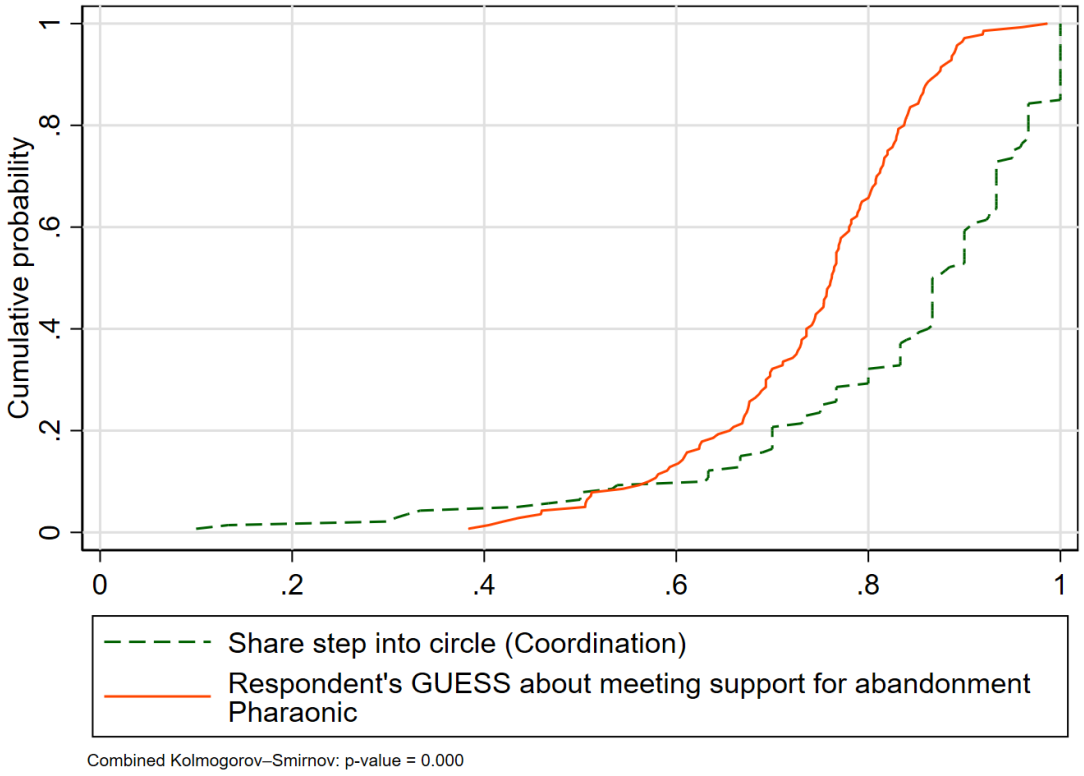


12 As for the voting cards, we did not record which participants stepped into the circle in Treatments 2 and 3. We only documented the share of participants in each meeting that stepped into the circle.

13 The two distributions are significantly different, according to the Kolmogorov-Smirnov test.

Figure 2 displays the CDF for the share of respondents who choose to step into the circle, represented by the dashed green line, and the respondents' average guess per meeting (the orange line – corresponding to the same distribution as in the previous figure). We see that the CDF for the share of respondents who publicly signal their support for abandonment of Pharaonic circumcision initially lies to the left of the CDF for the guess, while it eventually crosses it and lies to its right. This implies that, while in many communities the share of participants who publicly expressed support for abandoning Pharaonic exceeded participants' average guess, there were some communities where it fell short of it. In other words, the coordination exercise may have may have backfired in some of the communities. We return to this point in Section 4.2.4, where we explore mechanisms behind the treatment effects.

Figure 2: Beliefs about community support vs. Public signal of support for abandoning Pharaonic FGC: Evidence on coordination failure at baseline



3.3 Data

Our project comprised the collection of three rounds of data: a brief baseline that was collected at the start of the community meetings (November 2018 to May 2019), a midline survey 9 months after the intervention (January to June 2020), and an endline survey, one year and a half after the intervention, collected from March to May 2021. The follow-up surveys were conducted in all 141 communities on a sample of 4,152 individuals, equally split by gender (15 men and 15 women in each community), out of which

190 (4.6%) attrited at endline. Appendix Table A.2 shows that attrition was balanced across treatment groups. Individuals (and their households) were sampled from the list of 8,059 participants in the community meetings. As such, our sample is representative of the average participant in the meeting.[14] . The final sample of the endline survey also comprised information on respondents' daughters aged 0-18 years – 6,548 girls in total. All results reported in the paper refer to outcomes measured at endline, unless otherwise specified.

In the follow up surveys, individuals were asked questions about household composition and socioeconomic characteristics, expectations related to community norms, own cutting experience (female respondents only), and FGC status or intentions to cut for each of their daughters aged 0-18. We measured the prevalence of FGC in two ways: (i) direct questions where we ask respondents to report the FGC status of each daughter aged 0-18 and, if the respondent is female, their own status; and (ii) indirect revelation through a list experiment.[15]

Table 1: Descriptive statistics

	Mean	N
Respondents		
Female respondent (=1 female)	0.508	3962
Age	37.932	3962
Number of sons	2.886	3962
Number of daughters	2.825	3962
Education: No formal education	0.463	3962
Education: Incomplete primary	0.192	3962
Education: Secondary or above	0.104	3962
Education: Koranic study	0.121	3962
Daughters		
Pharaonic cut	0.034	6514
Sunna cut	0.480	6514
Not cut	0.469	6514
Age of cut	8.168	3461
Who decided to cut: Mother	0.829	3459

Table 1 presents some descriptive statistics for our respondents' and daugthers' sample. Half of the respondents are female by design. The average respondent is 38 years old and

14 Appendix Table A.1 displays the balance of baseline characteristics between individuals that were recruited for the survey and those who were not. While a few minor differences emerge, they are extremely small in magnitude: in no instance the normalized difference exceeds the value of 0.25 suggested by Imbens and Rubin (2015)

15 See Appendix B for details on the list experiment.

has about 3 sons and 3 daughters. The education level is very low: nearly half of the respondents have no formal education (46%), and one out of five has incomplete primary education. The lower panel of the table provides descriptive statistics about respondents' daughters aged 0-18 at the endline survey. On average 3.4% of them were Pharaonic cut, 48% of them were Sunna cut and 47% of them were uncut at the time of the survey. The average age of cutting was 8 years old. About 83% of the time, the decision related to FGC was taken by the girls' mother. Appendix Tables A.3, A.4, and A.5 present the balance table of individual characteristics across treatment arms. Almost all the variables are balanced across the four experimental conditions, confirming that the randomization was successful.[16]

4 Results

To estimate average treatment effects, we compare the outcomes of interest between individuals in treatment and control communities at follow-up. Thanks to the randomized design, these differences can be interpreted as causal. In particular, we estimate the following equation using ordinary least squares:

$$Y_{icd} = \alpha + \sum_{k=1}^{3} \beta_k T_{cd}^k + \lambda \mathbb{X}_{icd} + \theta_d + \varepsilon_{icd} \tag{2}$$

where Y_{icd} is the outcome of interest for respondent (or girl) i from community c in district d, T_{cd}^k are the three treatment dummies, θ_d are district fixed effects (randomization strata) and $\mathbb{X}_{icd}$ is a set of covariates used to improve precision.[17] We cluster standard errors at the community level, i.e., the unit of randomization.

In specification (2), the estimated β_1 represents the effect of revealing community preferences, β_2 the effect of the public declaration, and β_3 tests the effect of the combined intervention.

16 One variable that is unbalanced is respondents' age, which also causes differences in terms of number of children that respondents have. In our estimation, we control for these variables.

17 For outcomes at the respondent level, the covariates $\mathbb{X}_{icd}$ include: from baseline (meeting) data, (1) respondent's gender, (2) respondent's age, (3) respondent's number of sons, (4) respondent's number of daughters, and, from the endline survey, (5) respondent's education, (6) subclan dummies, (7) household socio-economic status, (8) respondent's marital status, (9) respondent's occupation dummies, (10) having a non-agricultural business (household level), (11) (if different from respondent) household head's age, (12) (if different from respondent) household head's education, (13) (if different from respondent) household head's occupation. For outcomes at the daughter level, in addition to the same household-level covariates, we control for the daughter's age and birth order. These covariates were all pre-specified as control variables at baseline, prior to any follow-up data collection.

4.1 Main Results

Table 2 presents the results of estimating specification (2) where the outcome of interest is the FGC status of respondent's daughters. The row "Mean in Control" at the bottom of Table 2 shows that in our sample of daughters aged 0-18, in the control group 4.3% of the girls are Pharaonic cut, 46% are Sunna cut, and 48% are still uncut at the time of the endline survey. The high proportion of uncut daughters is likely because many girls have not reached the cutting age yet.

In column 1, we find that correcting misperceptions by announcing the true support for the abandonment of Pharaonic FGC (T1) leads to a reduction in the prevalence of Pharaonic cut by 1.7 ppt (40% relative to the control mean). On the other hand, the coordination exercise (T2) did not have a significant impact, although the point estimate is negative. Similarly, the combined treatment (T3) had a negative but insignificant effect. Note that we cannot reject the null hypothesis that the effect of T1 is equal to that of T2 (p-value=0.248) or T3 (p-value=0.239).

In column 2, we estimate the impact of our interventions on the share of daughters who were cut with Sunna – the relatively less harmful type of FGC. We find that T1 led to a 6.2 ppt increase in the prevalence of Sunna (14% of control mean). While the effect of T2 is also positive, it is imprecisely estimated. The combined treatment has a marginally significant effect of 2.7 ppt (6% relative to the control mean). Finally, column 3 shows that T1 decreased the likelihood that respondents' daughters were uncut by the time of the survey by 4.7 ppt (a decrease of 9.7% with respect to the control group).

Overall, these results show that 1.5 years after the intervention was conducted, T1 led to a decrease in Pharaonic but an increase in Sunna FGC, and a reduction in the share of uncut girls. One potential explanation for the increase in Sunna circumcision could be that the information on community attitudes towards Pharaonic FGC may have been interpreted as more support for the intermediate norm (i.e., Sunna). Another possibility is that the move away from the Pharaonic norm occurs via an 'intermediate' step (i.e., the milder form of Sunna), as predicted by the theoretical framework of Gulesci et al. (2021).

Table 2: Treatment Effect on Daughters' FGC Rates

	(1) Pharaonic cut (yes=1)	(2) Sunna cut (yes=1)	(3) Uncut (yes=1)
T1 : Poll + Announcement	-0.017*	0.063***	-0.047***
	(0.009)	(0.018)	(0.015)
T2 : Poll + Coordination	-0.005	0.013	-0.004
	(0.009)	(0.016)	(0.015)
T3 : Poll + Announce. + Coord.	-0.006	0.027*	-0.013
	(0.009)	(0.016)	(0.014)
F pairwise:			
T1 vs T2	0.241	0.013	0.008
T1 vs T3	0.250	0.072	0.030
T2 vs T3	0.887	0.464	0.533
Mean in control	0.043	0.457	0.481
Observations	6,481	6,481	6,481
R^2	0.069	0.501	0.633

Notes: The sample includes respondent's daughters aged 18 or younger. The dependent variable in column 1 is a dummy variable =1 if the girl is Pharaonic cut; in column 2 is a dummy variable =1 if the girl is Sunna cut; and in column 3 is a dummy =1 if the girl is uncut. All regressions control for district fixed effects, daughter's age, daughter's birth order, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

During follow-up, respondents were asked about their plans for their daughters who were not circumcised at the time of the survey. We use this information to distinguish between girls who were uncut at the time of the survey and planned *never* to be cut, versus those who were uncut but planned to be cut at some point in the future. Given that nearly half of the girls in our sample are below the usual FGC age (of 9 years old), with this information we aim to shed light on the possible long-term effects of the interventions. Table 3 displays treatment effects on respondents' plans regarding their uncut daughters. In the control group, only 3.6% of girls aged 0-18 were uncut and planned not to be cut in the future (column 1). T1 led to a 3.2 ppt increase in this measure, corresponding to an 89% increase relative to the control group mean. While the effect of T2 is insignificant, the effect of the combined treatment (T3) is positive (1.6 ppt) and marginally significant, corresponding to a 44% increase relative to the control group.

Column 2 shows that 44% of the girls in the control group were uncut but planned to be cut at some point in the future. Relative to that, the girls in the T1 arm were 7.9 ppt (18%) less likely to be uncut but planned to get cut. Again, T2 has no significant effect, while T3 reduces this probability by 2.5 ppt (7%). This suggests that the negative effect

of T1 on girls' likelihood of being uncut at the time of the survey (Table 2, column 3) is mainly driven by girls who were *already* planned to be cut. In line with this, column 3 of Table 3 shows that T1 led to a decrease of 0.21 years (or 2.5 months) in the cutting age of girls cut or planned to be cut.

Table 3: Treatment Effect on Daughters' FGC Rates: Plans For Uncut Girls

	(1) Uncut, planned NOT to be cut (yes=1)	(2) Uncut, planned to be cut (yes=1)	(3) Cutting age (actual or planned)
T1 : Poll + Announcement	0.032***	-0.079***	-0.210*
	(0.010)	(0.017)	(0.107)
T2 : Poll + Coordination	0.007	-0.009	0.023
	(0.010)	(0.018)	(0.102)
T3 : Poll + Announce. + Coord.	0.016*	-0.029*	-0.075
	(0.009)	(0.017)	(0.100)
F pairwise:			
T1 vs T2	0.041	0.000	0.036
T1 vs T3	0.151	0.005	0.222
T2 vs T3	0.392	0.267	0.364
Mean in control	0.036	0.440	8.631
Observations	6,426	6,426	6,107
R^2	0.068	0.550	0.130

Notes: The sample includes respondent's daughters aged 18 or younger. The dependent variable in column 1 is a dummy variable =1 1 if the girl is uncut at the time of the survey and is NOT planned to be cut some time in the future; in column 2 is a dummy variable =1 if the girl is uncut and planned to be cut; and in column 3 is girls' cutting age, (actual, if cut, or planned, if uncut). All regressions control for district fixed effects, daughter's age, daughter's birth order, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

As an alternative way to evaluate the long-run effects of our interventions, we combine the information on respondents' daughters' current FGC status with the stated plans for daughters who were uncut at the time of the survey. Table 4 panel A shows that the announcement treatment (T1) decreased the likelihood that girls are or planned to be Pharaonic or Sunna circumcised by -1.8 ppt and -1.4 ppt respectively. Although these effects are statistically insignificant, together they lead to a 3.2 ppt increase in the likelihood of girls being uncut and planned to remain so. This is a significant effect, both statistically and economically, as it corresponds to a 89% reduction relative to control mean of 3.6%. On the other hand, the coordination treatment (T2) had no significant effect on any of the outcomes in Table 4 panel A. The combined treatment (T3) led to a marginally significant increase of 1.6 ppt (44%) in the likelihood that girls are and will remain uncut.

Table 4: Treatment effect on daughters' FGC rates: Potential future

Panel A: Actual daughters			
	(1) Pharaonic cut or plan to be (yes=1)	(2) Sunna cut or plan to be (yes=1)	(3) Uncut, planned NOT to be cut (yes=1)
T1 : Poll + Announcement	-0.018	-0.014	0.032***
	(0.011)	(0.017)	(0.010)
T2 : Poll + Coordination	-0.008	0.005	0.007
	(0.011)	(0.017)	(0.010)
T3 : Poll + Announce. + Coord.	-0.007	-0.004	0.016*
	(0.011)	(0.016)	(0.009)
F pairwise:			
T1 vs T2	0.390	0.275	0.041
T1 vs T3	0.311	0.541	0.151
T2 vs T3	0.981	0.604	0.392
Mean in control	0.054	0.889	0.036
Observations	6,428	6,428	6,428
R^2	0.061	0.037	0.068
Panel B: Hypothetical daughters			
	(1) Pharaonic cut (yes=1)	(2) Sunna cut (yes=1)	(3) Uncut (yes=1)
T1 : Poll + Announcement	-0.007	-0.110***	0.117***
	(0.007)	(0.024)	(0.022)
T2 : Poll + Coordination	0.007	-0.072**	0.064**
	(0.008)	(0.029)	(0.028)
T3 : Poll + Announce. + Coord.	-0.006	-0.077***	0.079***
	(0.006)	(0.025)	(0.024)
F pairwise:			
T1 vs T2	0.090	0.185	0.045
T1 vs T3	0.879	0.174	0.108
T2 vs T3	0.074	0.883	0.604
Mean in control	0.024	0.824	0.150
Observations	3,752	3,752	3,752
R^2	0.058	0.090	0.082

Notes: The sample of Panel A includes respondent's daughters aged 18 or younger. The sample of Panel B includes all respondents. In Panel A, the dependent variable in column 1 is a dummy variable =1 if the girl is Pharaonic cut or planned to be if uncut at the time of the survey; in column 2 is a dummy variable =1 if the girl is Sunna cut or planned to be if uncut at the time of the survey; and in column 3 is a dummy =1 if the girl is uncut and is NOT planned to be cut some time in the future. In Panel B, the dependent variable in column 1 is a dummy variable =1 if the respondent says that, if they have a daughter in the future, they will cut her and it will be Pharaonic; in column 2 is a dummy variable =1 if the respondents says that they will cut her and it will be Sunna; and in column 3 is a dummy variable =1 if the respondent says they will not cut her. All regressions control for district fixed effects, daughter's age (Panel A, only), daughter's birth order (Panel A only), respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

It is important to note that during the two years between the delivery of the treat-

ments and the endline data collection, not all the daughters in the sample reached the conventional cutting age. To further understand the effects of our treatment on respondents' *intentions* about FGC, we asked them to consider a scenario in which they may have another daughter in the future and whether they would cut her. If they said that they would cut her, we asked if they would choose Pharaonic or Sunna circumcision. Panel B of Table 4 reports the treatment effects on respondents' intentions. We find that all treatments increased the likelihood that respondents would choose not to cut a future hypothetical daughter by a magnitude ranging from 6.4 ppt (for T2) to 11.7 ppt (for T1). This is mainly driven by a decrease in hyppothetical Sunna circumcision.

Answers to questions of stated preferences involving hypothetical scenarios may suffer from hypothetical bias – i.e., a gap between real and stated preferences.[18] In particular, one may suspect that the large treatment effects in Table 4 may be due to experimenter demand.[19] We believe that the treatment effects in panel B of Table 4 indicate at least a shift in attitudes, if not intent. However, as further discussed in Section 4.4, we find evidence that these effects are not solely driven by experimenter demand. If anything, the discrepancy between control means and treatment effects in panels A and B could be because the former translates intent and the latter, attitudes. Benjamin et al. (2012) document that individuals usually make choices based on their subjective well-being, but this is not always the case if one also takes into account factors like family happiness or social status. It could be that our respondents *wish* future daughters are not circumcised, but under the given circumstances, this is not possible or practicable for current daughters[20] Even if these outcomes are hypothetical, it is interesting to note that the ordering of magnitudes of the treatment effects follows the ordering of panel A. The announcement treatment (Treatment 1) had larger effects than the coordination treatment (Treatment 2), and the combination of these two seems to have produced slightly smaller effects, even though this difference is not precisely estimated.

Taken together, the results presented in this section show that correcting misperceptions about the support for Pharaonic FGC decreased the likelihood that girls were Pharaonic cut but increased the likelihood of being Sunna cut at the time of the endline survey. Moreover, this treatment increased the likelihood that uncut girls may remain uncut in the long-run and respondents' intentions about hypothetical future daughters. On the other hand, the coordination treatment had no significant effect on FGC rates,

18 The literature documents that individuals may give untruthful answers to hypothetical questions out of concerns of social conformity and self-image, especially when choices have to do with morality (Johansson-Stenman and Svedsäter, 2012; Wuepper et al., 2019; Menapace and Raffaelli, 2020).

19 Note that our endline survey was delivered through an independent survey company, with no stated position regarding the issue of FGC. Even if this procedure does not eradicate all concerns of experimenter demand, it does make the purpose of our research less explicit to our respondents.

20 For about 8% of our sample of daughters, grandmothers (not parents) are the ones responsible for making circumcision decisions.

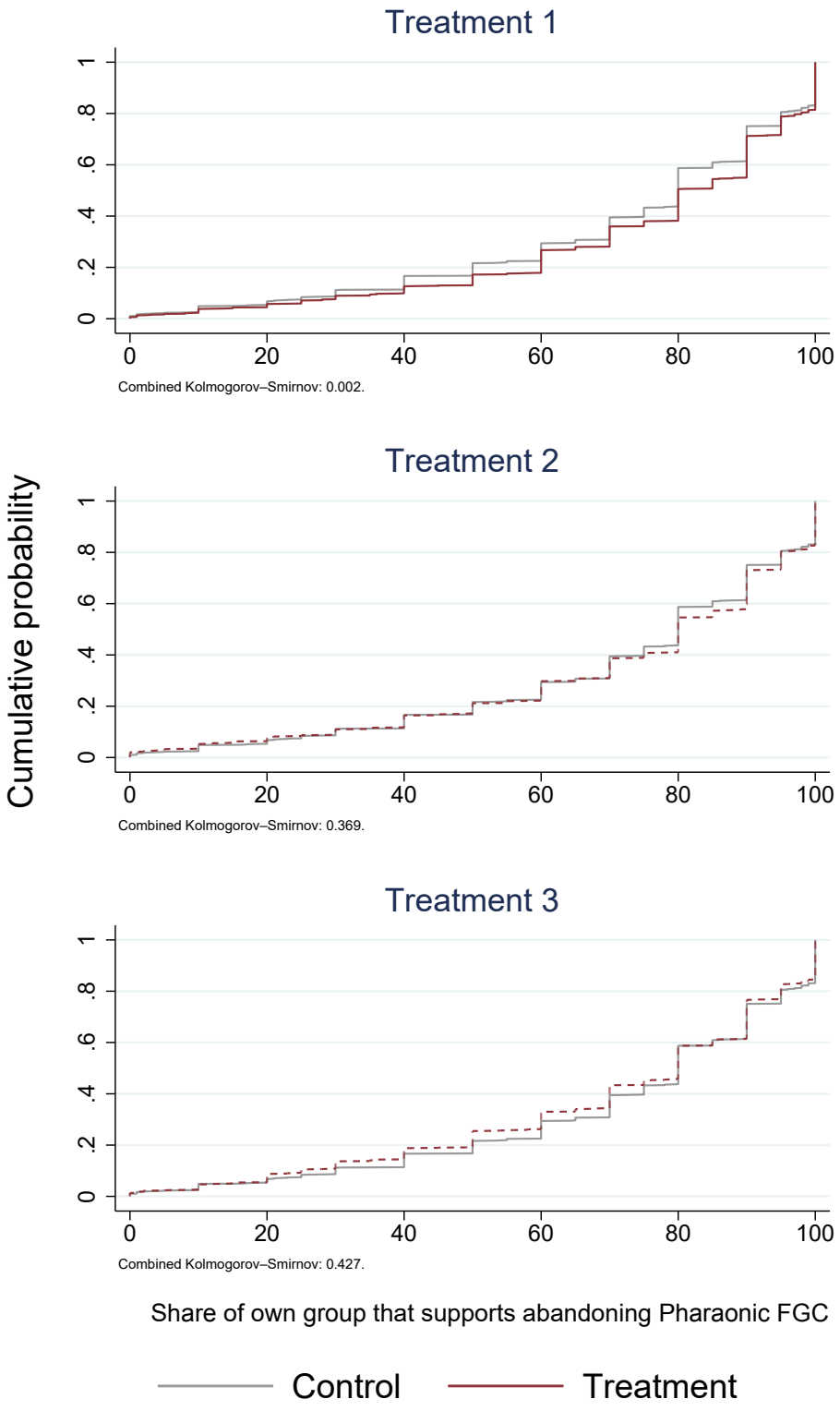
although it had some effects on respondents' intentions. The combination of both treatments presented smaller or null effects. In the next section, we explore the mechanisms driving these results.

4.2 Mechanisms

4.2.1 Information and belief updating

In Treatment 1, the announcement of the result of the vote should have corrected any misperceptions by revealing the "true" support for abandonment of Pharaonic FGC among the participants. This revelation should have brought individuals' beliefs on community support for abandoning Pharaonic closer to the true level. In order to assess this mechanism, we collected information on respondents' beliefs about the support for abandoning Pharaonic FGC within their communities. In particular, respondents were asked to report what percentage of women/men in their community supported the abandonment of Pharaonic circumcision. Figure 3 plots the distribution of respondents' second-order beliefs about support for abandoning Pharaonic FGC among their own gender group at midline (i.e., 9 months after the intervention was conducted), in the four experimental arms. We see that, in Treatment 1, relative to the control group, the CDF has shifted to the right (p-value=0.002). This suggests that the announcement caused participants to update their beliefs and made them believe a higher share of women/men within the community were in favor of abandoning Pharaonic FGC, at least 9 months after the intervention. On the other hand, the CDF for Treatment 2 and Treatment 3 lies more or less in the same position as the control group and Kolmogorov–Smirnov tests reveal that they are not statistically different from it. This implies that the coordination mechanism or the combined treatment did not lead to belief updating.

Figure 3: Beliefs about community support for abandoning FGC



To shed further light on the role of belief updating in driving the treatment effects, we analyze heterogeneity of the effects with respect to respondents' priors. If respondents update in a Bayesian way, the effect that the revelation of communities' true preferences

has on individuals should differ depending on the ex-ante expectations of each respondent regarding the share of community members in favor of abolishing infibulation. In a Bayesian updating framework, individuals whose prior was lower than the realized share should revise their expectations upwards, and vice versa. We test for this hypothesis, using information from the baseline survey where we elicited individuals' beliefs regarding the attitudes of others and by testing the heterogeneity of the impact along this dimension. In particular, we estimate:

$$Y_{icd} = \alpha + \sum_{k=1}^{3} \beta_1 T_{cd}^k + \sum_{k=1}^{3} \gamma_k T_{cd}^k (\overline{B_{cd}} - B_{icd0}) + \delta(\overline{B_{cd}} - B_{icd0}) + \lambda \mathbb{X}_{icd} + \theta_d + \varepsilon_{icd} \quad (3)$$

where $\overline{B_{cd}}$ is the level of support (for abolishing Pharaonic FGC) elicited in the community through the pool, and B_{icd0} is the initial prior (guess) of individual i. Under Bayesian updating, we expect $\gamma_1 < 0$ if Y corresponds to an indicator for being Pharaonic cut – the larger the gap between the true support and respondents' beliefs, the larger should be the magnitude of the effect of the announcement (Treatment 1) on likelihood that respondents choose Pharaonic circumcision for their daughters. We also estimate:

$$Y_{icd} = \alpha + \sum_{k=1}^{3} \beta_1 T_{cd}^k + \sum_{k=1}^{3} \gamma_k T_{cd}^k \mathbb{1}_{(\overline{B_{cd}} \geq B_{icd0})} + \delta \mathbb{1}_{(\overline{B_{cd}} \geq B_{icd0})} + \lambda \mathbb{X}_{icd} + \theta_d + \varepsilon_{icd} \quad (4)$$

where $\mathbb{1}_{(\overline{B_{cd}} \geq B_{icd0})}$ is a dummy variable equal to 1 if $\overline{B_{cd}} \geq B_{icd0}$ and 0 otherwise. This specification allows us to test if the effects are stronger among respondents who underestimated the community support (for abandoning Pharaonic FGC) at baseline.

Table 5 displays the results of estimating specifications (3) and (4) on respondents' daughters' FGC status at endline. In columns 1 and 2, we find that the estimate for γ_{10} has the predicted sign, although it is not statistically significant. For participants who underestimated the support for abandoning FGC, the announcement (Treatment 1) has a slightly larger effect on the likelihood their daughters are Pharaonic cut. In particular, the estimates in column 2 imply that the effect of Treatment 1 for respondents who did not underestimate support is -0.6 ppt and imprecisely estimated, while the effect for those who underestimated the support is -2.5 ppt and significant (p-value=0.018). We find no significant effect for the public declaration treatment (Treatment 2) or for the treatment arm combining both interventions (Treatment 3) for both the subsample of parents who underestimated support and those who did not. Panel A of figure XXX presents these results in a graphical format.

Table 5: Treatment Effect on Daughters' FGC Rates by Belief Wedge

	Pharaonic cut (yes=1)		Pharaonic cut or plan to be (yes=1)	
	(1)	(2)	(3)	(4)
T1 : Poll + Announcement	-0.016*	-0.006	-0.017	0.003
	(0.009)	(0.012)	(0.012)	(0.018)
T1 × Z	-0.004	-0.019	-0.032	-0.037*
	(0.025)	(0.014)	(0.032)	(0.021)
T2 : Poll + Coordination	-0.004	0.004	-0.007	0.005
	(0.010)	(0.015)	(0.012)	(0.017)
T2 × Z	-0.010	-0.014	-0.011	-0.020
	(0.032)	(0.018)	(0.034)	(0.022)
T3 : Poll + Announce. + Coord.	-0.006	0.002	-0.009	0.001
	(0.010)	(0.014)	(0.011)	(0.017)
T3 × Z	-0.013	-0.015	0.038	-0.014
	(0.029)	(0.018)	(0.045)	(0.021)
Z	-0.000	0.007	0.004	0.015
	(0.021)	(0.011)	(0.022)	(0.016)
Heterogeneity variable (Z)	$\bar{B} - B_{i0}$	$\mathbb{1}_{\bar{B} \geq B_{i0}}$	$\bar{B} - B_{i0}$	$\mathbb{1}_{\bar{B} \geq B_{i0}}$
Z mean	0.063	0.577	0.063	0.576
Dep. var mean in control	0.031	0.031	0.039	0.039
T1 + T1 × Z = 0	0.064	0.018	0.084	0.005
T2 + T2 × Z = 0	0.601	0.360	0.501	0.263
T3 + T3 × Z = 0	0.465	0.252	0.579	0.338
Observations	6,481	6,481	6,428	6,428
R^2	0.068	0.069	0.062	0.062

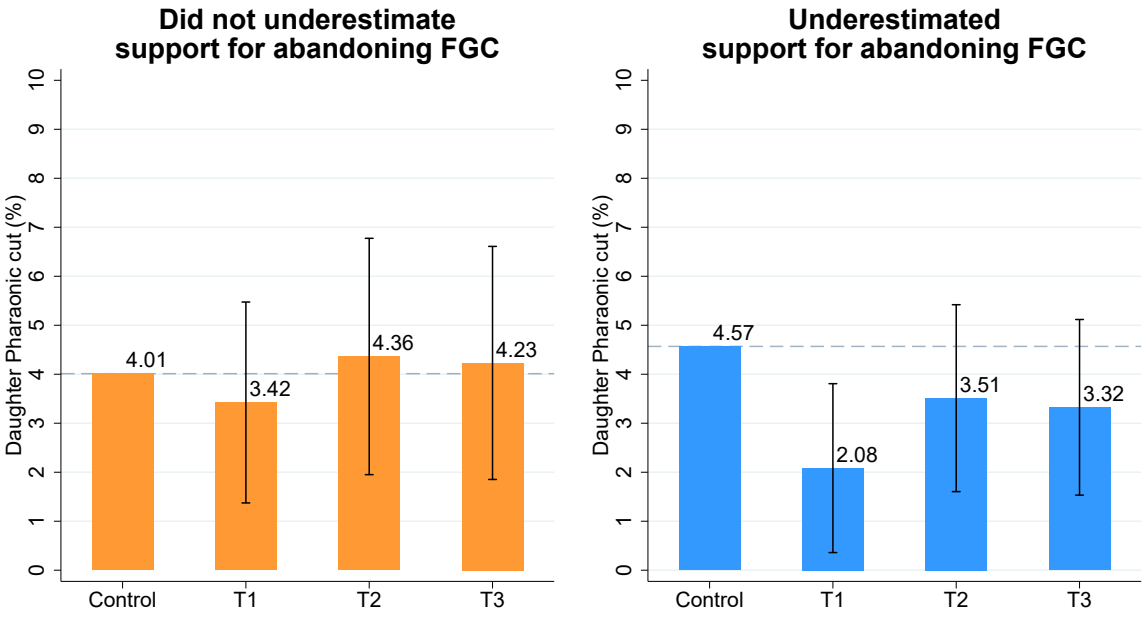
Notes: The sample includes respondent's daughters aged 18 or younger. The dependent variable in columns 1 and 2 is a dummy variable =1 if the girl is Pharaonic cut; in columns 3 and 4 is a dummy =1 if the girl is Pharaonic cut or uncut but planned to be Pharaonic cut. All regressions control for district fixed effects, daughter's age, daughter's birth order, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. T-tests in the bottom rows are evaluated at the mean value of Z at columns 1 and 3. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

In Table 5, we also present results on the heterogeneity of the treatment effects with respect to respondents' priors on the likelihood that girls are either Pharaonic cut at endline or planned to be Pharaonic cut at some point in the future. Similarly, we find that the treatment effects are larger for participants who underestimated the support for abandoning FGC. In particular, the estimates in column 4 imply that the effect of Treatment 1 for respondents who did not underestimate support is 0.03ppt and imprecisely estimated, while the effect for those who underestimated the support is -3.4ppt and significant (p-value=0.005). We find no significant effect for the public declaration

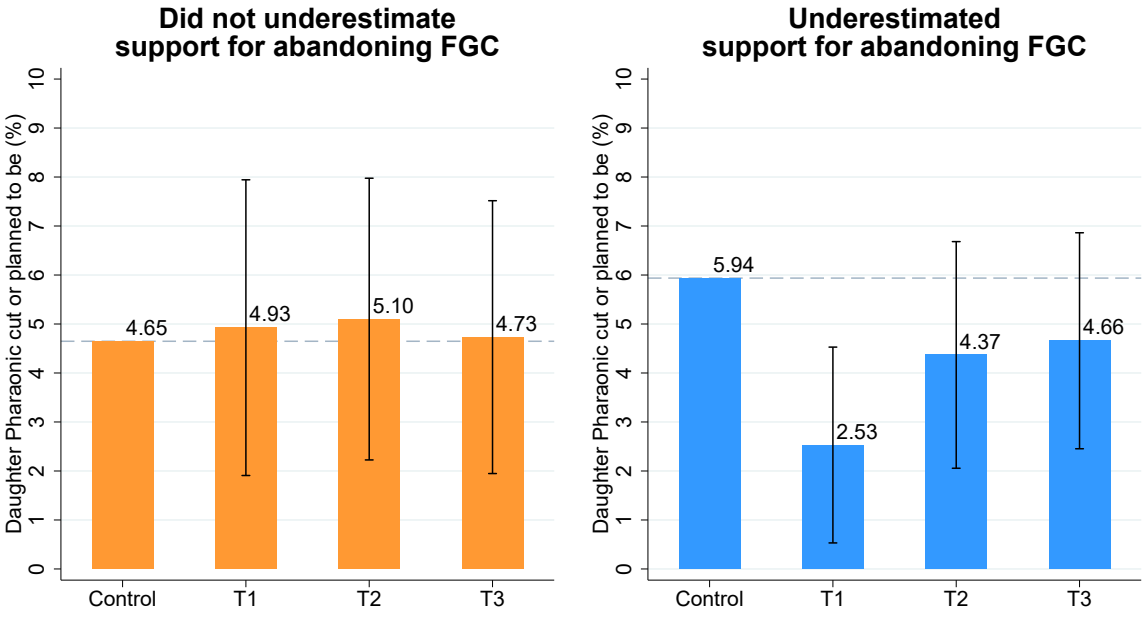
treatment (Treatment 2) or for the treatment arm combining both interventions (Treatment 3) for both the subsample of parents who underestimated support and those who did not. Panel B of Figure 4 displays these effects.

Figure 4: Treatment Effect on Daughters' FGC Rates by Belief Wedge

(A) Actual cutting status



(B) Actual cutting status and plans



Notes: The figure displays the estimated heterogeneous treatment effects of all treatments and their 90% confidence intervals (vertical axis). Estimates are presented in columns 2 and 4 of Table 5.

For completeness, we also estimate specifications (3) and (4) on likelihood that girls are Sunna cut or Uncut. Since the announcement treatment provided information about community members' support for Pharaonic circumcision, how this affects individuals' beliefs about alternative types of cut is unclear. On one hand, individuals who find out that support for Pharaonic is weaker than what they expected could interpret this as there being greater support for Sunna and switch to Sunna; or they could interpret it as support for not cutting and decide not to cut their daughters. Therefore, the prediction regarding the sign of the coefficient estimate for the interaction term (γ_1) is ambiguous. In line with this, Appendix Table A.6 shows that for Sunna cut, there is no significant heterogeneity in the treatment effects at the time of the survey. When we incorporate plans for uncut girls (columns 3-4) we find some evidence that effects are negative and marginally significant for individuals who found out the support for abandoning Pharaonic FGC to be lower than what they expected – but the interaction term is positive and significant in column 4, suggesting that in the long-run, people who underestimated the support at baseline were no more likely to switch to Sunna. Similarly, for uncut (columns 5-8) we do not find strong evidence of belief updating.

Overall, the results in this section provide evidence in line with belief updating as a mechanism driving the effects of the announcement treatment.

4.2.2 Indirect information on returns to FGC

Even though the announcement in Treatment 1 only provided information on the true support for Pharaonic FGC, it may indirectly convey information on the value of the practice. When individuals find out about the true support for Pharaonic being weaker than what they initially believed, they could ask around about the reasons for this. This may lead them to update their beliefs about the benefits of Pharaonic FGC. Table 6 provides evidence in line with this. In particular, we asked respondents if they would circumcise a hypothetical daughter in the future and if they say yes, then we asked why they would do so. In column 1, we find that Treatment 1 reduces the perceived religious value. Respondents were 8 ppt less likely to report that they would circumcise a future daughter because of religion. This is a 15% reduction relative to the control group. We do not find any significant effects of other treatments on this outcome. In column 2, we find that none of the treatments had a significant effect on expected likelihood to find a husband. This suggests that the effects were not driven by perceived returns on

the marriage market.[21] In column 3 of Table 6 we see that Treatment 1 had effects on perceived benefits in terms of preserving girls' virginity. In particular, respondents in Treatment 1 were 2.1 ppt less likely to report that they would cut their daughters because that would help preserve their virginity, which is a large reduction compared to the control group where only 4% of respondents said so. The effect of the other two treatments on these outcomes is imprecisely estimated and significantly lower than the effect of Treatment 1. These findings are in line with Treatment 1 leading to belief updating about the returns to Pharaonic FGC in terms of religious value or preserving girls' virginity.

Table 6: Treatment Effect on Reasons for FGC

	(1)	(2)	(3)	(4)
	Why would you cut a daughter in the future?			
	Religion	Find husb.	Virginity	Sex. feel.
T1 : Poll + Announcement	-0.079*	0.002	-0.021***	-0.004*
	(0.044)	(0.008)	(0.006)	(0.003)
T2 : Poll + Coordination	0.016	0.013	0.007	-0.004
	(0.048)	(0.008)	(0.008)	(0.003)
T3 : Poll + Announce. + Coord.	-0.003	0.003	-0.009	-0.003
	(0.045)	(0.009)	(0.007)	(0.003)
F pairwise:				
T1 vs T2	0.029	0.200	0.001	0.432
T1 vs T3	0.068	0.874	0.061	0.262
T2 vs T3	0.664	0.322	0.066	0.756
Mean in control	0.546	0.025	0.040	0.004
Observations	3,752	3,752	3,752	3,752
R^2	0.129	0.073	0.176	0.022

Notes: Observations are adult respondents. The dependent variable in column 1 is a dummy variable =1 if the respondent says that, if they had a daughter in the future, they would circumcise her because of religion; in column 2 is a dummy variable =1 if the respondent says that, if they had a daughter in the future, they would circumcise her to find a husband; in column 3 is a dummy variable =1 if the respondent says that, if they had a daughter in the future, they would circumcise her to keep her virginity; in column 4 is a dummy variable =1 if the respondent says that, if they had a daughter in the future, they would circumcise her to reduce sexual feelings. All regressions control for district fixed effects, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

21 As an alternative measure of perceived benefits in the marriage market, we asked respondents to state out of 100 unmarried men in their community, how many would be willing to marry a Sunna-cut girl or an Uncut girl. Appendix Table A.8 shows the treatment effects on these. On average, respondents in the control group believed that 73% of single men would be willing to marry a Sunna cut girl and 23% would marry an Uncut girl. We do not find any significant treatment effects on either of these outcomes. This reinforces the finding that none of the treatments led to a change in the perceived benefits of alternative forms of cutting in the marriage market.

The treatments could have also led to revelation of information about the negative health consequences of Pharaonic FGC or the (perceived to be) fewer negative health consequences of Sunna FGC. When individuals find that the true support for the practice is lower than what they initially expected, they could inquire about the reasons for this and find out about the health risks associated with it. To assess this, in Table 7, we estimate treatment effects on whether respondents believe Pharaonic or Sunna FGC to have any health consequences and the number of health risks they associate with Pharaonic or Sunna FGC. On average, 88% of the respondents in the control group believed there to be health consequences for Pharaonic FGC and on average they reported 2.4 distinct health risks associated with it, while for Sunna only 8% perceived there to be any health risks. We do not find any significant effects on these outcomes. Thus, we can rule out that Treatment 1 indirectly reveals information about the health consequences of the practice.

Table 7: Treatment Effect on Perceived Health Consequences of FGC

	(1) Any health consequences of Phar. cut	(2) Num. of health consequences of Phar. cut	(3) Any health consequences of Sunna cut	(4) Num. of health consequences of Sunna cut
T1 : Poll + Announcement	-0.021	0.093	0.001	0.049
	(0.031)	(0.157)	(0.018)	(0.041)
T2 : Poll + Coordination	-0.012	-0.042	0.009	0.028
	(0.025)	(0.139)	(0.022)	(0.047)
T3 : Poll + Announce. + Coord.	0.005	-0.025	0.007	0.039
	(0.029)	(0.171)	(0.018)	(0.040)
F pairwise:				
T1 vs T2	0.760	0.344	0.723	0.693
T1 vs T3	0.431	0.492	0.762	0.833
T2 vs T3	0.555	0.910	0.921	0.833
Mean in control	0.875	2.418	0.077	0.115
Observations	1,419	1,419	3,750	3,750
R^2	0.117	0.199	0.051	0.063

Notes: The dependent variable in column 1 is a dummy variable =1 if the respondent says that Pharaonic circumcision has any health consequences; in column 2 is the number of health consequences from Pharaonic circumcision that the respondent knows; in column 3 is a dummy variable =1 if the respondent says that Sunna circumcision has any health consequences; and in column 4 is the number of health consequences from Sunna circumcision that the respondent knows. Knowledge of health consequences was only asked if the type of circumcision was practiced in the community. All regressions control for district fixed effects, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

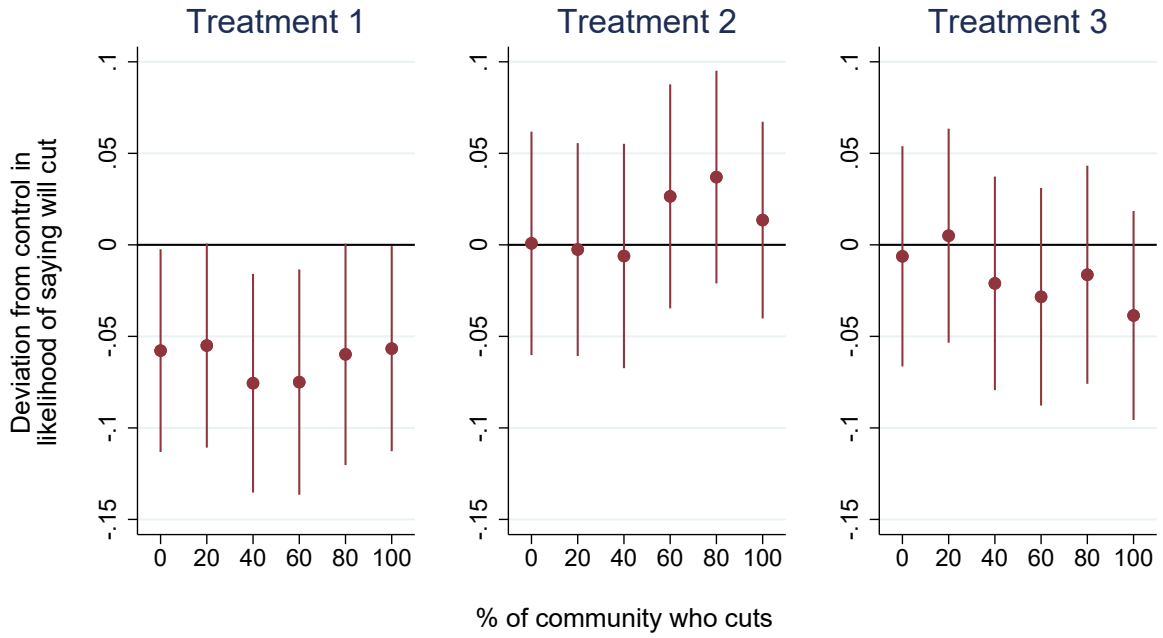
4.2.3 Social conformity

As described in Section 2, one explanation behind persistence of harmful norms is social conformity whereby individuals derive higher utility from taking a similar action as others. One mechanism through which Treatment 1 may have affected cutting rates could be through diminishing individuals' desire to conform with the majority and adopt the harmful behavior. First, correcting misperceptions about the support for the harmful norm may have shifted the perceived level of adoption of the norm and reduced the utility of conforming to it (i.e., in the utility function (1), a shift in $\sum_{j\neq i}^{N-1}\left(\mathbb{1}_{\omega_i=\omega_j}-\mathbb{1}_{\omega_i\neq\omega_j}\right)$ so less people are understood to conform with Pharaonic FGC). Second, it could also lead up to the understanding that, even when levels of adoption are high, support for the norm is weak and sanctions from not conforming are low (i.e., a shift in J). Lastly, it could also be that changes in perceived returns to the norm documented in Section 4.2.2 decreased the intrinsic utility of conforming to it (i.e., a shift in u_{ω_i}).

To shed light on this, we retrieved respondents' stated preferences for social conformity through their willingness to choose the harmful action as its prevalence within the community varies. In particular, we asked respondents if they would cut their daughters if x% of their community cut their daughters where x was either 100%, 80%, 60%, 40%, 20%, or 0%. These questions allow us to understand if treatments shifted J or u_{ω_i}. If treatments shifted J, we would expect them to have changed the slope of adoption across different levels of hypothetical community incidence. If treatments shifted u_h, we would expect them to affect equally the likelihood of choosing to cut across all different levels of hypothetical community incidence.

We find that neither Treatment 2 nor Treatment 3 shifted respondents' willingness to accept the norm, but Treatment 1 did (Figure 5). In particular, we find that respondents assigned to Treatment 1 were 6 to 8 ppt less likely to say they would cut their daughters if 40%, 60%, or all members in the community cut theirs (regression results are displayed in Appendix Table A.9). For the other cutoffs, coefficients are marginally insignificant. This suggests that correcting misperceptions led respondents to be less willing to conform with the majority in their community – they are less willing to choose the harmful action, even when the majority of their community does. On the other hand, the treatment effect remains similar as the prevalence of the harmful norm within the community varies. This suggests that Treatment 1 worked not so much by decreasing people's sensitivity to social pressure, but rather by conveying information on others' support (belief updating) and by reducing the perceived intrinsic value of the practice (as shown in Table 6 above).

Figure 5: "Suppose (...) individuals out of 100 in this community decided to circumcise their daughters, would you circumcise yours?"



Notes: The figure displays point estimates and 90% confidence intervals of treatment effects on respondents' likelihood to cut their daughters if x% of their community decide to circumcise their daughters. The corresponding regression results are displayed in Appendix Table A.9. The black line in each figure refers to the mean outcome in the control group.

4.2.4 Communication

All three treatments may have increased the level of communication among community members about FGC. In Treatment 1, the announcement of the true support for discontinuation of Pharaonic FGC may have started a dialogue among community members about the topic. In Treatment 2, the coordination exercise was designed to help individuals who favored the discontinuation of FGC but did not want to move unilaterally (for fear of social sanctions) to coordinate to their preferred equilibrium more easily, by lowering the cost of communication and coordination. In Table 8, we find that Treatment 2 significantly increased the likelihood of respondents talking to another household in the community by 7 ppt and of talking about FGC by 10 ppt on average. On the other hand, Treatment 1 and the combined treatment did not have any significant effects on communication – although we cannot reject the null of equality in effects of Treatment 2 and Treatment 3. On the intensive margin, individuals in Treatment 2 reported talking to fewer people – they were talking to 2 fewer people, which is a 16% reduction relative to the control mean of 12.7. These suggest that while Treatment 2 may have increased the likelihood of having a dialogue about FGC within the communities, it may have also led them to get less diverse perspectives. Moreover, it is not clear if more communication is

necessarily positive – they could talk to members with negative perspectives and reinforce conservative attitudes.

Table 8: Mechanisms: Communication

	(1) = 1 Talks to any HH	(2) Num. of HHs talks to	(3) =1 Talks to any HH about FGC	(4) Num. HHs talks to about FGC
T1 : Poll + Announcement	-0.008	-1.467	-0.006	-1.479
	(0.033)	(1.501)	(0.043)	(1.206)
T2 : Poll + Coordination	0.071**	-2.046*	0.098**	-1.532
	(0.034)	(1.225)	(0.047)	(1.125)
T3 : Poll + Announce. + Coord.	0.034	-0.460	0.055	-0.148
	(0.033)	(1.310)	(0.043)	(1.134)
F pairwise:				
T1 vs T2	0.020	0.690	0.008	0.964
T1 vs T3	0.190	0.512	0.076	0.263
T2 vs T3	0.246	0.238	0.267	0.249
Mean in control	0.849	12.720	0.739	8.630
Observations	3,955	3,955	3,955	3,955
R^2	0.099	0.175	0.099	0.201

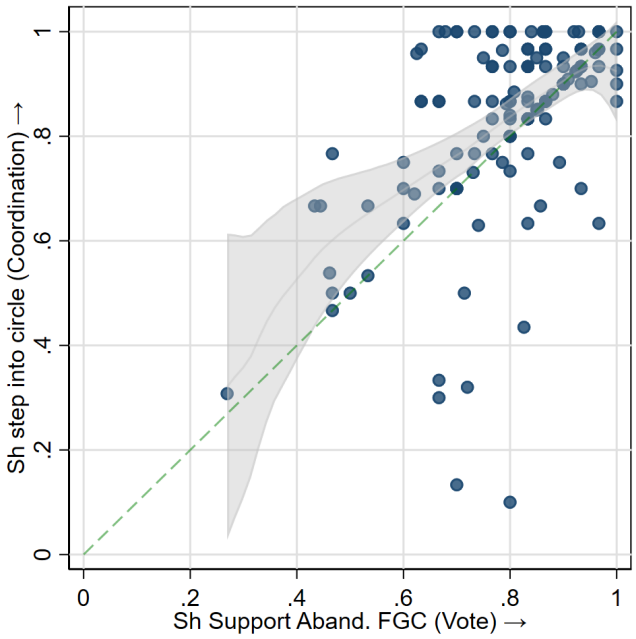
Notes: The dependent variable in column 1 is a dummy =1 if the respondents talk to any household in the community; in column 2 is the number of households in the community the respondent talks to (winsorized at the 99 level); in column 3 is a dummy =1 if the respondents talk to any household in the community about FGC; in column 4 is the number of households in the community the respondent talks to about FGC (winsorized at the 99 level). Observations are adult respondents. All regressions control for district fixed effects, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, a dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation categories. Robust standard errors are reported (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

4.2.5 Coordination

The results show that, in our setting, asking community members to make public statements to abandon FGC (which was done in both Treatment 2 and Treatment 3) is not effective in reducing cutting rates and may be counterproductive if combined with other interventions. One reason could be that public declarations may not perfectly translate private opinions and therefore participants interpret them differently. In most communities, the share who publicly signals support was greater or equal to the share who actually supports abandoning FGC, but not everywhere (see Figure 6). This could be because people may declare (or decide not to declare) their support due to social pressure or conformity concerns, which participants may take into account. For instance, participants that see fewer public declarations than their prior about community support may correct their priors downwards, or may believe that, even if private support to abandon the norm is high, people may not be willing to take public action against the norm. In case of the

combined arm (Treatment 3), in communities where the number of public declarations was different from the results of the poll disclosed in the meeting, participants might have also been driven to mistrust the announced results altogether.

Figure 6: Actual support vs. public signal of support for abandoning Pharaonic FGC (evidence on coordination failure at baseline)



The variation we observe could be because the decision to speak up against the harmful norm may hinge upon one's perception of what others in the meeting will choose to do (similarly to Chwe, 2000). In other words, differences in priors could explain the variation in the share of participants that stepped into the circle in meetings with similar voting patterns. In this sense, we would expect more participants to have stepped into the circle if they all believed there was large support for abandoning the practice, even if their private position was not so. Conversely, if participants held on average more negative priors, we would expect them to have been more hesitant to step into the circle, regardless of whether they privately supported abandoning the practice.

Table 9 displays the correlation between participants' average priors in meetings where the public declaration treatment was implemented (i.e., communities assigned to Treatment 2 or Treatment 3) and the share of participants who stepped into the circle. We observe that meetings with more positive priors on average had more participants stepping into the circle in order to publicly express their support for abandoning the practice. In particular, column 1 shows that a 1 ppt increase in participants' average second order beliefs is associated with a .35 ppt increase in the share of participants who stepped into the circle. Column 2 shows that in communities with high (i.e., above-median) priors about support for abandoning Pharaonic circumcision, 7 ppt more participants stepped

into the circle. As such, we would expect the coordination treatment to have been more effective in places where participants had more positive priors about the support for abandoning the practice.

Table 9: Mechanisms: Coordination and Average Priors in Meeting

	Sh. stepped inside circle	
	(1)	(2)
Z	0.345**	0.067***
	(0.132)	(0.023)
Num. participants	-0.001	-0.001
	(0.004)	(0.004)
Z	$\overline{B_{i0}}$	$\mathbb{1}_{\overline{B_{i0}} \geq med.}$
Z mean	0.733	0.514
Dep. var mean	0.829	0.829
Observations	140	140
R^2	0.436	0.437

Notes: The dependent variable in columns 1-2 is the share of participants that stepped into the circle (i.e., expressed public support against infibulation) in the baseline intervention. The independent variable (Z) is the meeting average prior about the share of participants that supported abandoning Pharaonic (column 1) and a dummy =1 if the meeting had an average meeting prior above the median of all meetings (column 2). Observations are meetings (2 per community). Only Treatment 2 and Treatment 3 meetings are reported. All regressions control for share of participants that privately support abandoning infibulation, number of participants in the meeting, and district fixed effects. Robust standard errors are reported (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

In line with this, in Table 10, we find that Treatment 2 decreased the prevalence of Pharaonic circumcision and the likelihood of girls being planned to be Pharaonic circumcised in meetings with more positive priors. Estimates in columns 1 and 3 indicate that Treatment 2, on average, did not have significant effects on daughters' likelihood of being Pharaonic or of being either Pharaonic or planned to be so. However, it seems that effects were more negative the larger the average prior in the meeting. The coefficient estimate for the interaction term is negative and large in columns 1 and 3 and significant in the latter. This suggests that effects were negative for participants that attended meetings with high average priors – in particular, negative and significant on whether daughters are Pharaonic or planned to be so. Estimates in columns 2 and 4 corroborate this observation and show that girls in meetings with above-median priors are 2.8 ppt less likely to be Pharaonic cut at endline, as also displayed in Figure 7, and 3.8 ppt less likely to be either Pharaonic cut or planned to be. Interaction terms for Treatment 3 generally have the same sign as those of Treatment 2 (with the exception of column 1)

but are imprecisely estimated.[22]

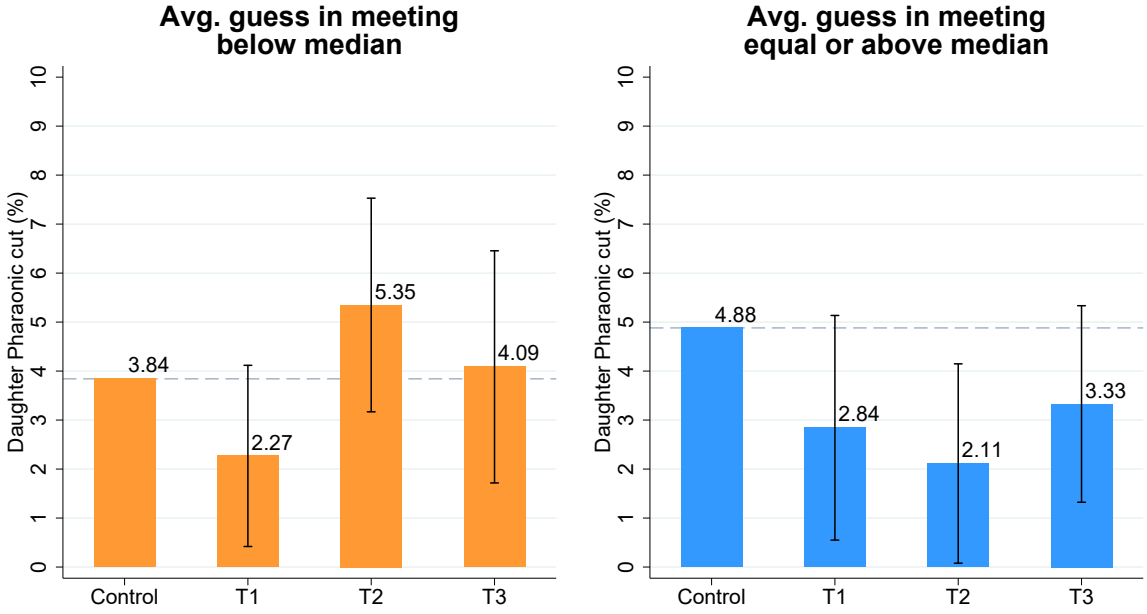
Table 10: Treatment Effect on Daughters' FGC Rates by Average Priors in Meeting

	Pharaonic cut (yes=1)		Pharaonic cut or plan to be (yes=1)	
	(1)	(2)	(3)	(4)
T1 : Poll + Announcement	-0.063	-0.016	-0.085	-0.025**
	(0.064)	(0.011)	(0.076)	(0.012)
T1 × Z	0.063	-0.005	0.090	0.009
	(0.087)	(0.018)	(0.107)	(0.022)
T2 : Poll + Coordination	0.084	0.015	0.155	0.020
	(0.076)	(0.013)	(0.096)	(0.015)
T2 × Z	-0.123	-0.043**	-0.225*	-0.058***
	(0.102)	(0.018)	(0.129)	(0.022)
T3 : Poll + Announce. + Coord.	-0.035	0.002	0.025	0.005
	(0.070)	(0.014)	(0.086)	(0.017)
T3 × Z	0.040	-0.018	-0.043	-0.025
	(0.093)	(0.019)	(0.117)	(0.024)
Z	-0.047	0.015	-0.020	0.015
	(0.083)	(0.015)	(0.103)	(0.019)
Heterogeneity variable (Z)	$\overline{B_{i0}}$	$\mathbb{1}_{\overline{B_{i0}} \geq med.}$	$\overline{B_{i0}}$	$\mathbb{1}_{\overline{B_{i0}} \geq med.}$
Z mean	0.737	0.524	0.737	0.523
Dep. var mean in control	0.031	0.031	0.039	0.039
T1 + T1 × Z = 0	0.053	0.143	0.075	0.374
T2 + T2 × Z = 0	0.481	0.026	0.371	0.014
T3 + T3 × Z = 0	0.507	0.201	0.531	0.208
Observations	6,481	6,481	6,428	6,428
R^2	0.070	0.070	0.065	0.065

Notes: The sample includes respondent's daughters aged 18 or younger. The dependent variable in columns 1 and 2 is a dummy variable =1 if the girl is Pharaonic cut; in columns 3 and 4 is a dummy =1 if the girl is Pharaonic cut or uncut but planned to be Pharaonic cut. The independent variable (Z) is the meeting average prior about the share of participants that supported abandoning Pharaonic (columns 1 and 3) and a dummy =1 if the meeting had an average meeting prior above the median of all meetings (columns 2 and 4). T-tests in the bottom rows are evaluated at the mean value of Z at columns 1 and 3. All regressions control for district fixed effects, daughter's age, daughter's birth order, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

22 Estimates for Sunna circumcision and uncut are displayed in Appendix Table A.7. Since both treatments targeted Pharaonic circumcision, the prediction for the treatment effects on Sunna and uncut are less clear. We find that in communities with higher average priors (about support for abandoning Pharaonic circumcision) the coordination exercise led to an increase in likelihood of being Sunna-cut and a decrease in uncut by the endline survey. The effects become insignificant when we incorporate the plans for uncut girls.

Figure 7: Treatment Effect on Daughters' FGC Rates by Average Priors in Meeting (actual cuts)



Notes: The figure displays estimated heterogeneous treatment effects of all treatments and their 90% confidence intervals (vertical axis). Estimates are presented in column 2 of Table 10.

4.3 Robustness checks

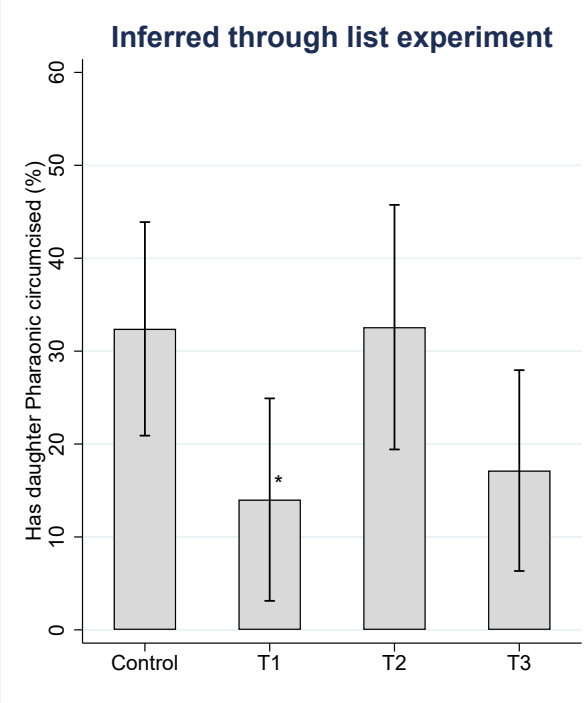
Given that FGC could be perceived as a sensitive topic, we also analyzed respondents' behavior using list experiments (also known as the "item count technique")[23] . We implemented a 4-item list, further described in Appendix B, where half of our sample was randomly selected to receive a fifth item containing the potentially sensitive statement "You have a daughter who has been Pharaonic cut". All participants were asked to indicate the number of statements that were true to them. The true share of respondents that have a Pharaonic circumcised daughter can be inferred by comparing the average number of correct statements in the groups with and without the potentially sensitive statement. It is important to highlight that the sensitive item did not specify the age of the daughter. Thus, the levels of circumcision rates may differ from the rates reported in the previous tables, where the sample was restricted only to daughters below age 18.

Using the item list method, we infer that on average 32.4% of respondents in the control group have a Pharaonic cut daughter (Figure 8). This share is smaller for those in Treatment 1 (14.0%) – the difference is significant at the 10-percent level. While participants in Treatment 2 are as likely as those in the control group to have a Pharaonic cut daughter, those in Treatment 3 are notably less likely (17.1%), although the differ-

23 In another project in Sierra Leone, Corno and La Ferrara (2023) have found that self-reported measures of FGC are consistent with objective measures collected through health checks. In Sierra Leone, just like Somalia, FGC is nearly universal and there is no domestic legislation banning it.

ence between T3 and control is not precisely estimated. These results are qualitatively consistent with the treatment effects in Table 2. They suggest that the announcement treatment (Treatment 1) is the most effective intervention in reducing the incidence of the most harmful form of FGC, whereas no effect is observed in the public declaration treatment (Treatment 2) and effects are generally smaller in the treatment arm combining both interventions (Treatment 3). Point estimates are displayed in Appendix Table B.5.

Figure 8: Share of respondents that have a Pharaonic cut daughter, inferred through list experiment and stated



Notes: The figure displays the percentage of respondents that have Pharaonic circumcised daughters (of any age) in each treatment arm and 90% confidence intervals. The four bars display the estimates inferred using the item-count technique. Bar 1 displays the mean difference in the item-count lists of groups with and without the sensitive item in the control group. Bars 2-4 display the linear combination between bar 1 and the coefficients estimated in Appendix Table B.5. Stars indicate significant differences with the control group (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

5 Conclusions

Understanding why harmful practices such as female genital cutting persist is necessary to design effective policies that can reduce their prevalence. In this paper, we explored if the prevalence of infibulation in Somali communities can be explained by two different mechanisms: pluralistic ignorance and coordination failures. We evaluated three different interventions to test these hypotheses. The first revealed the true share of support to the harmful norm, the second offered a coordination device to abandon it, and the third did both.

We find that on average community members overestimate the community support for infibulation and correcting this misperception is effective in reducing its incidence.

This finding is robust to using measures of incidence of infibulation inferred through list experiments. Revealing the true support appears to affect cutting rates through belief updating as well as lower perceived returns of cutting in terms of religious value and chastity. The coordination device does not significantly affect infibulation on average. It increases communication among households in the extensive margin, but it leads to a decrease in communication in the intensive margin. This may have led participants to get less diverse views on FGC, but it could also be that more communication about FGC is not positive and reinforces conservative attitudes. Nevertheless, in communities with more positive priors where participants are more likely to express public support to the norm, the coordination device seems to work and decrease the incidence of infibulation. Finally, combining both interventions seems to produce smaller effects on average. We believe this is likely because the combination introduces extra noise on the signal we send to respondents about the support to the norm.

Our findings suggest that a simple, low-cost intervention that corrects individuals' misperceptions about their community's support for infibulation can be an effective way of reducing its prevalence. On the other hand, designing and implementing interventions that effectively address coordination failures requires a better understanding of the interplay between beliefs and coordination.

References

Abdalla, R. H. D. (1982). *Sisters in Affication: Circumcision and Infibulation of Women in Africa*. Connecticut: Lawrence Hill and Co.

Adam, T., H. Bathija, D. Bishai, Y. Bonnenfant, M. Darwish, D. Huntington, and E. Johansen (2010). Estimating the obstetric costs of female genital mutilation in six african countries. *Bulletin of the World Health Organization 88*, 281.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics 128*(2), 469–530.

Ambrus, A. and E. Field (2008). Early marriage, age of menarche and female schooling attainment in bangladesh. *Journal of Political Economy 116*(5), 881–930.

Anderson, S. (2007). The economics of dowry and brideprice. *Journal of Economic Perspectives 21*(4), 151–174.

Anderson, S. and C. Bidner (2015). Property rights over marital transfers. *The Quarterly Journal of Economics 130*(3), 1421–1484.

Andreoni, J., N. Nikiforakis, and S. Siegenthaler (2021). Predicting social tipping and norm change in controlled experiments. *Proceedings of the National Academy of Sciences 118*(16), e2014893118.

Apffelstaedt, A., J. Freundt, and C. Oslislo (2022). Social norms and elections: How elected rules can make behavior (in) appropriate. *Journal of Economic Behavior & Organization 196*, 148–177.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2020). Bride price and female education. *Journal of Political Economy 28*(2), 591–641.

Barman-Adhikari, A., A. Al Tayyib, S. Begun, E. Bowen, and E. Rice (2017). Descriptive and injunctive network norms associated with nonmedical use of prescription drugs among homeless youth. *Addictive Behaviors 64*, 70–77.

Becker, A. (2018). On the economic origins of female genital cutting. Working paper.

Benjamin, D. J., O. Heffetz, M. S. Kimball, and A. Rees-Jones (2012). What do you think would make you happier? what do you think you would choose? *American Economic Review 102*(5), 2083–2110.

Bhalotra, S., A. Chakravarty, and S. Gulesci (2020). The price of gold: Dowry and death in india. *Journal of Development Economics 143*, 102413.

Bicchieri, C. and A. Marini (2016). Female genital cutting: Fundamentals, social expectations and change. MPRA working paper No. 72927.

Brener, L., C. von Hippel, R. Horwitz, and J. Hamwood (2015). The impact of pluralistic ignorance on the provision of health care for people who inject drugs. *Journal of Health Psychology 20*(9), 1240–1249.

Brock, W. A. and S. N. Durlauf (2001). Discrete choice with social interactions. *Review of Economic Studies 68*(2), 235–260.

Brock, W. A. and S. N. Durlauf (2002). A Multinomial-Choice Model of Neighborhood Effects. *The American Economic Review (P&P) 92*(2), 298–303.

Brock, W. A. and S. N. Durlauf (2005). *Multinomial Choice with Social Interactions*, Volume III of *The Economy as an Evolving Complex System*, Chapter 8. Oxford University Press.

Bursztyn, L., A. W. Cappelen, B. Tungodden, A. Voena, and D. H. Yanagizawa-Drott (2023, March). How are gender norms perceived? Working Paper 31049, National Bureau of Economic Research.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American economic review 110*(10), 2997–3029.

Caiado, C. C., W. A. Brock, R. A. Bentley, and M. J. O'Brien (2016). Fitness landscapes among many options under social influence. *Journal of Theoretical Biology 405*, 5–16. Advances in Modelling Biological Evolution: Linking Mathematical Theories with Empirical Realities.

Camilotti, G. (2016). Interventions to stop female genital cutting and the evolution of the custom: Evidence on age at cutting in senegal. *Journal of African Economies 25*(1), 133–158.

Chwe, M. S.-Y. (2000). Communication and coordination in social networks. *The Review of Economic Studies 67*(1), 1–16.

Cislaghi, B., E. K. Denny, M. Cisse, P. Gueye, B. Shrestha, P. N. Shrestha, G. Ferguson, C. Hughes, and C. J. Clark (2019). Changing Social Norms: the Importance of "Organized Diffusion" for Scaling Up Community Health Promotion and Women Empowerment Interventions. *Prevention Science 20*(6), 936–946.

Corno, L. and E. La Ferrara (2023). Norms Replacement. A Field Experiment on ending Female Genital Cutting. mimeo.

Corno, L., E. La Ferrara, and A. Voena (2020). Female genital cutting and the slave trade. Working paper.

De Souza, L. and T. Schmader (2022). The misjudgment of men: Does pluralistic ignorance inhibit allyship? *Journal of Personality and Social Psychology 122*(2), 265.

Diop, N. J., M. Faye, A. Moreau, J. Cabral, H. Benga, F. Cissé, B. Mané, I. Baumgarten, and M. Melching (2004). The tostan program. evaluation of a community based education program in senegal.

Duffy, J. and J. Lafky (2021). Social conformity under evolving private preferences. *Games and Economic Behavior 128*, 104–124.

Durlauf, S. N. and Y. M. Ioannides (2010). Social interactions. *Annual Review of Economics 2*(1), 451–478.

Efferson, C., S. Vogt, A. Elhadi, H. Ahmed, and E. Fehr (2015). Female genital cutting is not a social coordination norm. *Science 349*(6255), 1446–1447.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society*, 1047–1071.

Fernandez, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics 1*(1), 146–177.

Fernandez, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *Quarterly Journal of Economics 119*(4), 1249–1299.

Fernández-Duque, M. (2022). The probability of pluralistic ignorance. *Journal of Economic Theory 202*, 105449.

Gele, A. A., M. Sagbakken, and B. Kumar (2015). Is female circumcision evolving or dissolving in norway? A qualitative study on attitudes toward the practice among young Somalis in the Oslo area. *International journal of women's health*, 933–943.

Giuliano, P. (2007). Living arrangements in western europe: Does cultural origin matter? *Journal of the European Economic Association 5*(5), 927–952.

Gulesci, S., S. Jindani, E. La Ferrara, D. Smerdon, M. Sulaiman, and H. Young (2021). A stepping stone approach to understanding harmful norms.

Hagemeister, F. (2022). Populism and propagation of far-right extremism. *European Journal of Political Economy 72*, 102116.

Jaldesa, G. W., I. Askew, C. Njue, and M. Wanjiru (2005). Female genital cutting among the somali of kenya and management of its complications. *Population Council/FRONTIERS/USAID*.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics 7*, 63–88.

Jensen, R. and R. Thornton (2003). Early female marriage in the developing world. *Gender & Development 11*(2), 9–19.

Johansson-Stenman, O. and H. Svedsäter (2012). Self-image and valuation of moral goods: Stated versus actual willingness to pay. *Journal of Economic Behavior & Organization 84*(3), 879–891.

Karaffa, K. M. and J. M. Koch (2016). Stigma, pluralistic ignorance, and attitudes toward seeking mental health services among police officers. *Criminal justice and behavior 43*(6), 759–777.

Katz, D., F. H. Allport, and M. B. Jenness (1931). Students' attitudes; a report of the syracuse university reaction study.

Kreindler, G. E. and H. P. Young (2013). Fast convergence in evolutionary equilibrium selection. *Games and Economic Behavior 80*, 39–67.

Mackie, G. (1996). Ending footbinding and infibulation: A convention account. *American Sociological Review 61*(6), 999–1017.

Menapace, L. and R. Raffaelli (2020). Unraveling hypothetical bias in discrete choice experiments. *Journal of Economic Behavior & Organization 176*, 416–430.

Miyajima, T. and H. Yamaguchi (2017). I want to but i won't: Pluralistic ignorance inhibits intentions to take paternity leave in japan. *Frontiers in Psychology 8*, 1508.

Novak, L. (2020). Persistent norms and tipping points: The case of female genital cutting. *Journal of Economic Behavior & Organization 177*, 433 – 474.

O'Gorman, H. J. (1975). Pluralistic ignorance and white estimates of white support for racial segregation. *Public Opinion Quarterly 39*(3), 313–330.

Platteau, J.-P., G. Camilotti, and E. Auriol (2018). Eradicating women-hurting customs: What role for social engineering? In S. Anderson, L. Beaman, and J.-P. Platteau (Eds.), *Towards gender equity in development*. Oxford: Oxford University Press.

Prentice, D. A. and D. T. Miller (1993). Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm. *Journal of personality and social psychology 64*(2), 243.

Sandstrom, M., H. Makover, and M. Bartini (2013). Social context of bullying: Do misperceptions of group norms influence children's responses to witnessed episodes? *Social Influence 8*(2-3), 196–215.

Smerdon, D., T. Offerman, and U. Gneezy (2020). Everybody's doing it: On the persistence of bad social norms. *Experimental Economics 23*, 392–420.

Soetevent, A. R. and P. Kooreman (2007). A discrete-choice model with social interactions: With an application to high school teen behavior. *Journal of Applied Econometrics 22*(3), 599–624.

Somalia National Bureau of Statistics (2020). The Somali Health and Demographic Survey 2020. https://somalia.unfpa.org/en/publications/somali-health-and-demographic-survey-2020.

Spisma, H. L., P. G. Chen, A. Ofori-Atta, U. O. Ilozumba, K. Karfo, and E. H. Bradley (2012). Female genital cutting: current practices and beliefs in western africa. *Bulletin of the World Health Organization 90*, 120–127F.

UNICEF (2008). Long-term evaluation of the Tostan programme in Senegal: Kolda, Thiès and Fatick regions. www.unicef.org/evaldatabase/index_59605.html.

UNICEF (2013). Female genital mutilation/cutting: A statistical overview and exploration of the dynamics of change. https://www.unicef.org/publications/index_69875.html.

UNICEF (2016). Female genital mutilation/cutting: A global concern. https://data.unicef.org/wp-content/uploads/2016/04/FGMC-2016-brochure_250.pdf.

UNICEF (2022). Eliminating female genital mutilation: Unicef 2021 annual report to the u.s. department of state. https://www.unicef.org/documents/eliminating-female-genital-mutilation.

Van Grootel, S., C. Van Laar, L. Meeussen, T. Schmader, and S. Sczesny (2018). Uncovering pluralistic ignorance to change men's communal self-descriptions, attitudes, and behavioral intentions. *Frontiers in psychology 9*, 1344.

Vogt, S., N. Zaid, H. Ahmed, E. Fehr, and C. Efferson (2016). Female circumcision: A critical appraisal. *Nature 538*, 506–509.

WHO (2023). Female genital mutilation. `https://www.who.int/news-room/fact-sheets/detail/female-genital-mutilation`. Retrieved on March 13, 2023.

Wuepper, D., A. Clemm, and P. Wree (2019). The preference for sustainable coffee and a new approach for dealing with hypothetical bias. *Journal of Economic Behavior & Organization 158*, 475–486.

Yoder, P. S., S. Wang, and E. Johansen (2013). Estimates of female genital mutilation/cutting in 27 african countries and yemen. *Studies in Family Planning 44*(2), 189–204.

Young, H. P. (2015). The evolution of social norms. *Annual Review of Economics 7*(1), 359–387.

Yu, H. (2023). Social stigma as a barrier to HIV testing: Evidence from a randomized experiment in Mozambique. *Journal of Development Economics 161*, 103035.

Online Appendix
Changing Harmful Norms through Information and Coordination. Experimental Evidence from Somalia

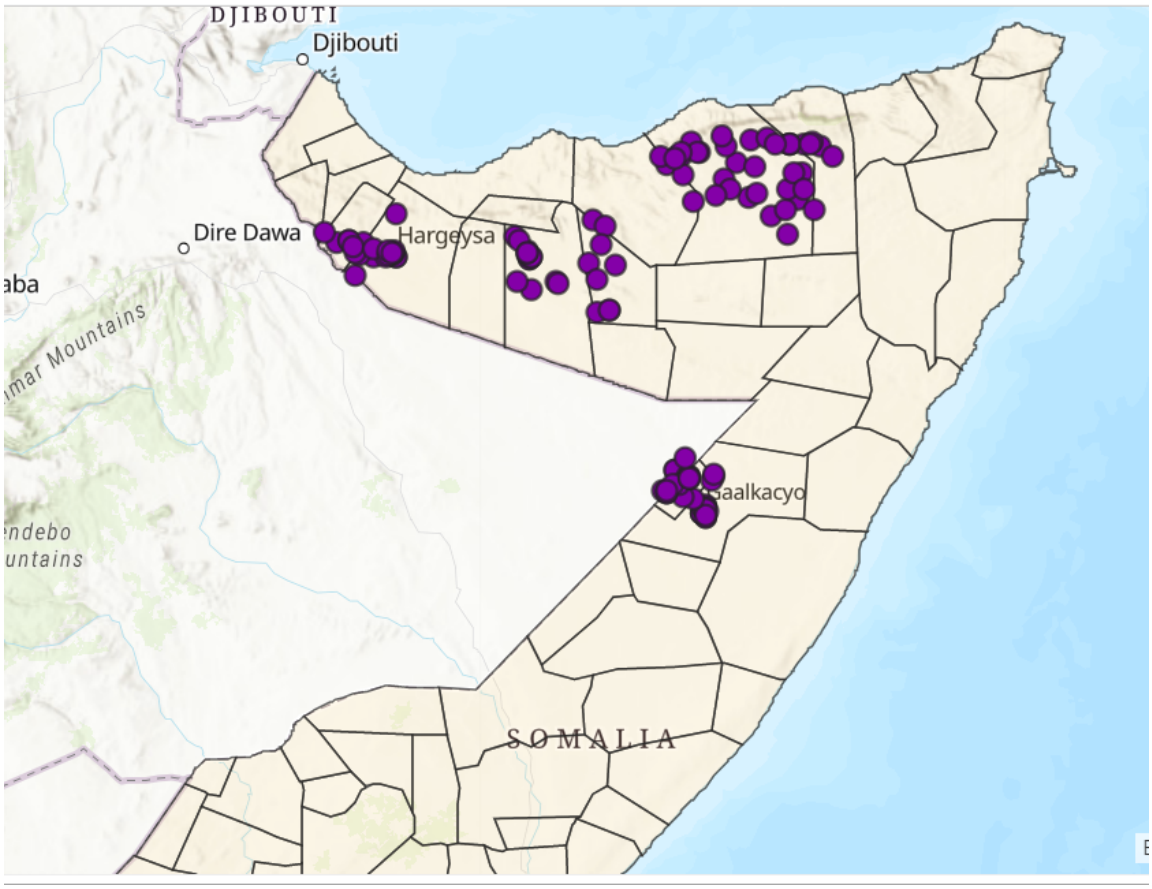
16 September 2023

Pedro Ferreira[1]
Selim Gulesci[2]
Eliana La Ferrara[3]
David Smerdon[4]
Munshi Sulaiman[5]

1 Harvard University and LEAP; pedro_ferreira@hks.harvard.edu.
2 Trinity College Dublin; gulescis@tcd.ie.
3 Harvard University, CEPR, NBER and LEAP; elaferrara@hks.harvard.edu.
4 University of Queensland; d.smerdon@uq.edu.au.
5 BRAC; munshi.sulaiman@brac.net.

A Additional figures and tables

Figure A.1: Location of the Study



Notes: The map shows the location of communities included in the study, along with district boundaries of Somalia.

Table A.1: Experimental design

Group	**Intervention**	**Number of communities**
Control	Poll only	35
Treatment 1	Poll + Announcement of the result of the poll	36
Treatment 2	Poll + Coordination exercise	34
Treatment 3	Poll + Announcement of the result of the poll + Coordination exercise	36

Table A.1: Sample Selection

	(1) **Recruited** Mean/SD	(2) **Non-Recruited** Mean/SD	(3) **P-value**	(4) **Norm diff.**
Demographic characteristics				
Female	0.508	0.512	0.728	0.006
	(0.500)	(0.500)		
Age	37.881	36.998	0.012**	-0.044
	(14.111)	(14.460)		
Number of sons	2.864	2.681	0.001***	-0.057
	(2.275)	(2.283)		
Number of daughters	2.813	2.597	0.000***	-0.068
	(2.267)	(2.226)		
Baseline guesses characteristics				
Guess of sh. of own group that supp. abandon Phar.	0.741	0.718	0.000***	-0.065
	(0.241)	(0.250)		
Guess of sh. of other group that supp. abandon Phar.	0.655	0.657	0.787	0.005
	(0.262)	(0.259)		
Guess of sh. of grandmothers that supp. abandon Phar.	0.589	0.592	0.640	0.008
	(0.294)	(0.293)		
Guess of sh. of grandfathers that supp. abandon Phar.	0.605	0.609	0.556	0.009
	(0.287)	(0.283)		
District of origin				
Hargeisa	0.210	0.210	0.976	0.001
	(0.407)	(0.408)		
Burao	0.166	0.174	0.282	0.014
	(0.372)	(0.379)		
Erigavo	0.142	0.148	0.256	0.012
	(0.349)	(0.356)		
Badhan	0.146	0.116	0.001***	-0.063
	(0.354)	(0.321)		
Southern Galkayo	0.173	0.185	0.470	0.021
	(0.379)	(0.388)		
Galdogob	0.162	0.167	0.459	0.010
	(0.368)	(0.373)		
Observations	4152	3907		

Table A.2: Attrition

	=1 Attrition		
	(1)	(2)	(3)
T1 : Poll + Announcement	0.003	0.002	0.003
	(0.015)	(0.014)	(0.014)
T2 : Poll + Coordination	0.010	0.006	0.007
	(0.014)	(0.013)	(0.013)
T3 : Poll + Announce. + Coord.	0.003	-0.000	0.001
	(0.014)	(0.013)	(0.013)
Mean in control	0.955	0.958	0.958
District FE	No	Yes	Yes
Baseline controls	No	No	Yes
Observations	4,152	4,140	4,140
R^2	0.000	0.005	0.008

Notes: Observations are adult respondents recruited for the follow-up. Baseline characteristics include: gender, age, and number of sons and daughters. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table A.3: Balance Table (A): Baseline characteristics and demographic characteristics

	(1) **Control** (C)		(2) **Announcement** (T1)		(3) **Coordination** (T2)		(4) **Announc. + Coord.** (T3)		(5) **P-value** (T1-C)	(6) **P-value** (T2-C)	(7) **P-value** (T3-C)	(8) **Norm diff.** (T1-C)	(9) **Norm diff.** (T2-C)	(10) **Norm diff.** (T3-C)
	N	Mean/SD	N	Mean/SD	N	Mean/SD	N	Mean/SD						
Baseline guesses characteristics														
Guess of sh. of own group that supp. abandon Phar.	978	0.726	994	0.758	968	0.726	1021	0.757	0.145	0.941	0.132	0.094	0.002	0.035
		(0.247)		(0.235)		(0.241)		(0.234)						
Guess of sh. of other group that supp. abandon Phar.	978	0.642	994	0.683	968	0.644	1021	0.660	0.059*	0.939	0.443	0.112	0.006	0.022
		(0.262)		(0.255)		(0.268)		(0.259)						
Guess of sh. of grandmothers that supp. abandon Phar.	978	0.571	994	0.605	968	0.588	1021	0.591	0.190	0.455	0.406	0.082	0.040	0.025
		(0.290)		(0.290)		(0.295)		(0.300)						
Guess of sh. of grandfathers that supp. abandon Phar.	978	0.585	994	0.626	968	0.603	1021	0.608	0.089*	0.412	0.381	0.102	0.045	0.027
		(0.289)		(0.286)		(0.285)		(0.287)						
Demographic characteristics														
Female	978	0.521	994	0.502	968	0.507	1021	0.503	0.298	0.349	0.289	-0.028	-0.020	-0.021
		(0.500)		(0.500)		(0.500)		(0.500)						
Age	978	39.603	994	35.757	968	37.871	1021	38.479	0.000***	0.106	0.252	-0.196	-0.086	-0.073
		(14.068)		(13.651)		(14.336)		(14.106)						
Number of sons	978	3.129	994	2.731	968	2.814	1021	2.867	0.009***	0.041**	0.066*	-0.124	-0.099	-0.093
		(2.261)		(2.274)		(2.253)		(2.319)						
Number of daughters	978	3.081	994	2.625	968	2.699	1021	2.895	0.005***	0.021**	0.216	-0.143	-0.124	-0.066
		(2.242)		(2.268)		(2.121)		(2.393)						
Ever married	973	0.952	986	0.902	964	0.911	1016	0.915	0.009***	0.110	0.036**	-0.136	-0.115	-0.037
		(0.215)		(0.298)		(0.285)		(0.278)						
Age at marriage	923	21.311	888	20.598	874	21.286	925	21.546	0.041**	0.937	0.578	-0.078	-0.003	0.028
		(6.883)		(5.971)		(6.960)		(6.812)						
Married a relative	923	0.384	879	0.367	881	0.393	928	0.386	0.641	0.692	0.976	-0.023	0.013	0.003
		(0.487)		(0.482)		(0.489)		(0.487)						
Woman brought any asset into marriage	492	0.081	469	0.077	461	0.063	494	0.083	0.911	0.399	0.984	-0.012	-0.050	0.004
		(0.274)		(0.266)		(0.243)		(0.276)						
Subclan: Majarteen (Darood)	978	0.303	994	0.295	968	0.235	1020	0.292	0.168	0.455	0.246	-0.012	-0.109	-0.015
		(0.460)		(0.456)		(0.424)		(0.455)						
Subclan: Others (Darood)	978	0.198	994	0.146	968	0.200	1020	0.143	0.168	0.822	0.126	-0.099	0.004	-0.100
		(0.399)		(0.353)		(0.401)		(0.350)						
Subclan: Garxajis (Dir)	978	0.189	994	0.174	968	0.126	1020	0.168	0.639	0.167	0.515	-0.028	-0.123	-0.038
		(0.392)		(0.379)		(0.332)		(0.374)						
Subclan: Habar Awal (Dir)	978	0.090	994	0.185	968	0.146	1020	0.125	0.136	0.416	0.431	0.197	0.123	0.076
		(0.286)		(0.389)		(0.353)		(0.330)						
Subclan: Habar Jeclo (Dir)	978	0.125	994	0.109	968	0.132	1020	0.134	0.389	0.687	0.543	-0.035	0.016	0.019
		(0.331)		(0.311)		(0.339)		(0.341)						
Subclan: Other (Dir)	978	0.054	994	0.040	968	0.073	1020	0.105	0.527	0.745	0.187	-0.047	0.055	0.128
		(0.227)		(0.197)		(0.261)		(0.307)						
Subclan: Others	978	0.041	994	0.051	968	0.088	1020	0.033	0.605	0.161	0.983	0.035	0.136	-0.028
		(0.198)		(0.221)		(0.283)		(0.180)						

Table A.4: Balance Table (B): Socioeconomic characteristics

	(1) **Control** (C)		(2) **Announcement** (T1)		(3) **Coordination** (T2)		(4) **Announc. + Coord.** (T3)		(5) **P-value** (T1-C)	(6) **P-value** (T2-C)	(7) **P-value** (T3-C)	(8) **Norm diff.** (T1-C)	(9) **Norm diff.** (T2-C)	(10) **Norm diff.** (T3-C)
	N	Mean/SD	N	Mean/SD	N	Mean/SD	N	Mean/SD						
Socioeconomic characteristics														
Asset index	978	0.212	994	0.179	968	0.065	1021	0.161	0.769	0.342	0.872	-0.018	-0.083	-0.037
		(1.330)		(1.257)		(1.163)		(1.167)						
HH runs non-agricultural business	978	0.093	994	0.124	968	0.088	1021	0.171	0.189	0.673	0.001***	0.070	-0.013	0.155
		(0.291)		(0.329)		(0.283)		(0.377)						
Education: No formal education	940	0.457	968	0.487	949	0.497	988	0.468	0.606	0.558	0.820	0.041	0.057	0.012
		(0.498)		(0.500)		(0.500)		(0.499)						
Education: Incomplete primary	940	0.200	968	0.208	949	0.176	988	0.207	0.722	0.317	0.710	0.013	-0.043	0.012
		(0.400)		(0.406)		(0.381)		(0.406)						
Education: Primary, incomplete secondary	940	0.079	968	0.110	949	0.084	988	0.095	0.089*	0.655	0.309	0.075	0.014	0.040
		(0.269)		(0.312)		(0.278)		(0.294)						
Education: Secondary or above	940	0.106	968	0.105	949	0.122	988	0.097	0.918	0.648	0.802	-0.002	0.035	-0.021
		(0.308)		(0.307)		(0.328)		(0.296)						
Education: Koranic study	940	0.157	968	0.091	949	0.120	988	0.133	0.012**	0.317	0.599	-0.143	-0.076	-0.048
		(0.364)		(0.288)		(0.325)		(0.339)						
Occupation: day-laborer	973	0.326	988	0.306	966	0.329	1017	0.291	0.892	0.544	0.645	-0.031	0.005	-0.049
		(0.469)		(0.461)		(0.470)		(0.454)						
Occupation: farmer	973	0.122	988	0.139	966	0.122	1017	0.110	0.890	0.786	0.555	0.034	0.000	-0.026
		(0.328)		(0.346)		(0.328)		(0.313)						
Occupation: self-employed (non-agricultural)	973	0.067	988	0.068	966	0.048	1017	0.080	0.993	0.185	0.490	0.003	-0.058	0.034
		(0.250)		(0.252)		(0.213)		(0.271)						
Occupation: salaried work	973	0.072	988	0.068	966	0.081	1017	0.073	0.738	0.594	0.876	-0.011	0.023	0.002
		(0.259)		(0.252)		(0.273)		(0.260)						
Occupation: housework	973	0.197	988	0.165	966	0.149	1017	0.199	0.329	0.119	0.785	-0.059	-0.090	0.002
		(0.398)		(0.371)		(0.356)		(0.399)						
Occupation: student	973	0.030	988	0.040	966	0.034	1017	0.033	0.320	0.631	0.580	0.041	0.018	0.015
		(0.170)		(0.197)		(0.182)		(0.180)						
Occupation: other	973	0.186	988	0.215	966	0.237	1017	0.214	0.441	0.125	0.316	0.050	0.088	0.047
		(0.389)		(0.411)		(0.425)		(0.411)						

Table A.5: Balance Table (C): Household head characteristics and district of origin

	(1)		(2)		(3)		(4)		(5)	(6)	(7)	(8)	(9)	(10)
	Control		**Announcement**		**Coordination**		**Announc. + Coord.**		**P-value**	**P-value**	**P-value**	**Norm diff.**	**Norm diff.**	**Norm diff.**
	(C)		(T1)		(T2)		(T3)		(T1-C)	(T2-C)	(T3-C)	(T1-C)	(T2-C)	(T3-C)
	N	Mean/SD	N	Mean/SD	N	Mean/SD	N	Mean/SD						
Household head characteristics														
Age	977	43.551	990	41.251	965	42.331	1017	43.334	0.019**	0.201	0.727	-0.116	-0.061	-0.014
		(13.762)		(14.353)		(14.673)		(14.124)						
Education: No formal education	949	0.445	962	0.500	943	0.490	989	0.466	0.217	0.451	0.929	0.078	0.064	0.025
		(0.497)		(0.500)		(0.500)		(0.499)						
Education: Incomplete primary	949	0.195	962	0.192	943	0.177	989	0.200	0.955	0.468	0.757	-0.005	-0.032	0.009
		(0.396)		(0.394)		(0.382)		(0.400)						
Education: Primary, incomplete secondary	949	0.082	962	0.106	943	0.086	989	0.092	0.183	0.757	0.523	0.058	0.009	0.024
		(0.275)		(0.308)		(0.280)		(0.289)						
Education: Secondary or above	949	0.118	962	0.112	943	0.123	989	0.104	0.803	0.880	0.676	-0.013	0.011	-0.030
		(0.323)		(0.316)		(0.329)		(0.306)						
Education: Koranic study	949	0.160	962	0.089	943	0.124	989	0.138	0.007***	0.331	0.683	-0.152	-0.073	-0.043
		(0.367)		(0.285)		(0.330)		(0.345)						
Occupation: day-laborer	977	0.392	990	0.362	965	0.370	1018	0.361	0.763	0.997	0.824	-0.044	-0.032	-0.041
		(0.488)		(0.481)		(0.483)		(0.480)						
Occupation: farmer	977	0.123	990	0.135	965	0.124	1018	0.119	0.955	0.826	0.707	0.026	0.003	-0.008
		(0.328)		(0.342)		(0.330)		(0.324)						
Occupation: self-employed (non-agricultural)	977	0.059	990	0.063	965	0.049	1018	0.072	0.891	0.409	0.508	0.010	-0.033	0.035
		(0.236)		(0.242)		(0.215)		(0.258)						
Occupation: salaried work	977	0.089	990	0.072	965	0.098	1018	0.079	0.246	0.709	0.498	-0.045	0.023	-0.026
		(0.285)		(0.258)		(0.298)		(0.269)						
Occupation: housework	977	0.151	990	0.130	965	0.122	1018	0.147	0.436	0.254	0.663	-0.043	-0.060	-0.008
		(0.359)		(0.337)		(0.328)		(0.355)						
Occupation: student	977	0.021	990	0.025	965	0.022	1018	0.022	0.605	0.878	0.798	0.018	0.001	0.001
		(0.145)		(0.157)		(0.146)		(0.145)						
Occupation: other	977	0.164	990	0.213	965	0.215	1018	0.201	0.120	0.105	0.200	0.089	0.092	0.065
		(0.370)		(0.410)		(0.411)		(0.401)						
Daughters' characteristics														
Daughter's age	1670	9.425	1623	9.205	1592	9.234	1661	9.709	0.376	0.390	0.275	-0.032	-0.027	0.048
		(4.971)		(4.886)		(5.015)		(5.012)						
Daughter's birth order	1670	1.939	1623	1.962	1592	1.913	1661	1.907	0.775	0.551	0.525	0.015	-0.017	-0.018
		(1.078)		(1.163)		(1.050)		(1.088)						

Table A.6: Treatment Effect on Daughters' FGC Rates (Sunna and Uncut) by Belief Wedge

	Sunna cut (yes=1)		Sunna cut or plan to be (yes=1)		Uncut (yes=1)		Uncut, planned NOT to be cut (yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1 : Poll + Announcement	0.068***	0.063**	-0.018	-0.044*	-0.053***	-0.062***	0.033***	0.038**
	(0.018)	(0.027)	(0.017)	(0.026)	(0.015)	(0.021)	(0.011)	(0.016)
T1 × Z	-0.082	-0.001	0.082	0.054*	0.098*	0.027	-0.043	-0.012
	(0.062)	(0.034)	(0.054)	(0.032)	(0.054)	(0.029)	(0.034)	(0.019)
T2 : Poll + Coordination	0.017	0.016	0.003	-0.010	-0.011	-0.019	0.005	0.006
	(0.017)	(0.026)	(0.018)	(0.024)	(0.016)	(0.021)	(0.010)	(0.013)
T2 × Z	-0.045	-0.004	0.019	0.025	0.081*	0.025	0.016	0.001
	(0.062)	(0.035)	(0.054)	(0.031)	(0.047)	(0.026)	(0.029)	(0.016)
T3 : Poll + Announce. + Coord.	0.029*	0.013	-0.005	-0.019	-0.018	-0.015	0.017*	0.021
	(0.017)	(0.026)	(0.016)	(0.024)	(0.014)	(0.021)	(0.009)	(0.014)
T3 × Z	-0.009	0.025	-0.000	0.026	0.041	0.000	-0.027	-0.009
	(0.059)	(0.036)	(0.059)	(0.030)	(0.048)	(0.031)	(0.029)	(0.016)
Z	0.037	0.001	-0.017	-0.013	-0.053	-0.015	-0.003	-0.006
	(0.044)	(0.026)	(0.037)	(0.023)	(0.032)	(0.021)	(0.017)	(0.010)
Heterogeneity variable (Z)	$\bar{B} - B_{i0}$	$\mathbb{1}_{\bar{B} \geq B_{i0}}$	$\bar{B} - B_{i0}$	$\mathbb{1}_{\bar{B} \geq B_{i0}}$	$\bar{B} - B_{i0}$	$\mathbb{1}_{\bar{B} \geq B_{i0}}$	$\bar{B} - B_{i0}$	$\mathbb{1}_{\bar{B} \geq B_{i0}}$
Z mean	0.063	0.577	0.063	0.576	0.063	0.577	0.063	0.576
Dep. var mean in control	0.487	0.487	0.886	0.886	0.464	0.464	0.054	0.054
T1 + T1 × Z = 0	0.001	0.007	0.453	0.639	0.002	0.075	0.002	0.024
T2 + T2 × Z = 0	0.379	0.596	0.794	0.522	0.729	0.745	0.556	0.593
T3 + T3 × Z = 0	0.083	0.084	0.756	0.734	0.304	0.508	0.083	0.231
Observations	6,481	6,481	6,428	6,428	6,481	6,481	6,428	6,428
R^2	0.501	0.501	0.038	0.038	0.633	0.633	0.069	0.069

Notes: The sample includes respondent's daughters aged 18 or younger. The dependent variable in columns 1 and 2 is a dummy variable =1 if the girl is Sunna cut; in columns 3 and 4 is a dummy =1 if the girl is Sunna cut or uncut but planned to be Sunna cut; in columns 5 and 6 is a dummy variable =1 if the girl is uncut; in columns 7 and 8 is a dummy =1 if the girl is uncut and planned not to be cut. All regressions control for district fixed effects, daughter's age, daughter's birth order, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table A.7: Treatment Effect on Daughters' FGC Rates (Sunna and Uncut) by Average Priors in Meeting

	Sunna cut (yes=1)		Sunna cut or plan to be (yes=1)		Uncut (yes=1)		Uncut, planned NOT to be cut (yes=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1 : Poll + Announcement	0.065	0.079***	0.128	0.016	-0.010	-0.057**	-0.057	0.018
	(0.112)	(0.025)	(0.113)	(0.022)	(0.091)	(0.023)	(0.082)	(0.016)
T1 × Z	0.003	-0.022	-0.193	-0.053*	-0.055	0.013	0.121	0.027
	(0.150)	(0.036)	(0.152)	(0.031)	(0.123)	(0.032)	(0.109)	(0.023)
T2 : Poll + Coordination	-0.282***	-0.021	-0.113	-0.004	0.232***	0.012	-0.006	-0.009
	(0.104)	(0.022)	(0.143)	(0.024)	(0.074)	(0.020)	(0.075)	(0.014)
T2 × Z	0.407***	0.071**	0.162	0.019	-0.324***	-0.033	0.018	0.033
	(0.143)	(0.031)	(0.193)	(0.033)	(0.105)	(0.028)	(0.102)	(0.021)
T3 : Poll + Announce. + Coord.	-0.097	0.003	0.074	0.008	0.119	0.005	-0.094	0.001
	(0.119)	(0.024)	(0.130)	(0.025)	(0.100)	(0.019)	(0.074)	(0.014)
T3 × Z	0.171	0.047	-0.107	-0.023	-0.183	-0.036	0.150	0.030
	(0.160)	(0.033)	(0.175)	(0.033)	(0.138)	(0.028)	(0.100)	(0.019)
Z	-0.120	-0.028	0.105	0.017	0.151	0.018	-0.088	-0.022*
	(0.117)	(0.023)	(0.127)	(0.021)	(0.095)	(0.021)	(0.066)	(0.012)
Heterogeneity variable (Z)	$\overline{B_{i0}}$	$\mathbb{1}_{\overline{B_{i0}} \geq med.}$	$\overline{B_{i0}}$	$\mathbb{1}_{\overline{B_{i0}} \geq med.}$	$\overline{B_{i0}}$	$\mathbb{1}_{\overline{B_{i0}} \geq med.}$	$\overline{B_{i0}}$	$\mathbb{1}_{\overline{B_{i0}} \geq med.}$
Z mean	0.737	0.524	0.737	0.523	0.737	0.524	0.737	0.523
Dep. var mean in control	0.487	0.487	0.886	0.886	0.464	0.464	0.054	0.054
T1 + T1 × Z = 0	0.000	0.027	0.429	0.113	0.001	0.036	0.002	0.002
T2 + T2 × Z = 0	0.283	0.026	0.719	0.514	0.681	0.315	0.520	0.142
T3 + T3 × Z = 0	0.092	0.020	0.761	0.469	0.315	0.133	0.055	0.011
Observations	6,481	6,481	6,428	6,428	6,481	6,481	6,428	6,428
R^2	0.502	0.502	0.039	0.039	0.633	0.633	0.069	0.069

Notes: The sample includes respondent's daughters aged 18 or younger. The dependent variable in columns 1 and 2 is a dummy variable =1 if the girl is Sunna cut; in columns 3 and 4 is a dummy =1 if the girl is Sunna cut or uncut but planned to be Sunna cut; in columns 5 and 6 is a dummy variable =1 if the girl is uncut; in columns 7 and 8 is a dummy =1 if the girl is uncut and planned not to be cut. The independent variable (Z) is the meeting average prior about the share of participants that supported abandoning Pharaonic (columns 1 and 3) and a dummy =1 if the meeting had an average meeting prior above the median of all meetings (columns 2 and 4). All regressions control for district fixed effects, daughter's age, daughter's birth order, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table A.8: Mechanism: Perceived Returns on Marriage Market

	(1) Out of 100 unmarried men, how many would marry girl not Phar but Sunna	(2) Out of 100 unmarried men, how many would marry girl not Phar but uncut
T1 : Poll + Announcement	0.929	-0.989
	(2.345)	(2.558)
T2 : Poll + Coordination	4.046	2.064
	(2.461)	(2.507)
T3 : Poll + Announce. + Coord.	2.087	-0.682
	(2.266)	(2.528)
F pairwise:		
T1 vs T2	0.161	0.117
T1 vs T3	0.566	0.875
T2 vs T3	0.372	0.167
Mean in control	72.844	22.705
Observations	3,728	3,325
R^2	0.155	0.102

Notes: Observations are adult respondents. The dependent variable in column 1 is the respondent's estimate of the number (out of 100) of unmarried men that would be willing to marry a girl that is not Pharaonic cut, but Sunna; in column 2 is the respondent's estimate of the number (out of 100) of unmarried men that would be willing to marry a girl that is not Pharaonic cut and uncut. All regressions control for district fixed effects, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table A.9: Social conformity

	(1) =1 would cut daug. if 100 out of 100 cut theirs	(2) =1 would cut daug. if 80 out of 100 cut theirs	(3) =1 would cut daug. if 60 out of 100 cut theirs	(4) =1 would cut daug. if 40 out of 100 cut theirs	(5) =1 would cut daug. if 20 out of 100 cut theirs	(6) =1 would cut daug. if no one cut theirs
T1 : Poll + Announcement	-0.057*	-0.060	-0.075**	-0.076**	-0.055	-0.058*
	(0.034)	(0.037)	(0.037)	(0.036)	(0.034)	(0.033)
T2 : Poll + Coordination	0.014	0.037	0.026	-0.006	-0.003	0.001
	(0.032)	(0.035)	(0.037)	(0.037)	(0.035)	(0.037)
T3 : Poll + Announce. + Coord.	-0.039	-0.016	-0.028	-0.021	0.005	-0.006
	(0.034)	(0.036)	(0.036)	(0.035)	(0.035)	(0.036)
F pairwise:						
T1 vs T2	0.032	0.006	0.007	0.061	0.132	0.086
T1 vs T3	0.613	0.244	0.215	0.140	0.101	0.134
T2 vs T3	0.129	0.126	0.135	0.683	0.836	0.848
Mean in control	0.793	0.753	0.740	0.695	0.633	0.579
Observations	3,962	3,962	3,962	3,962	3,962	3,962
R^2	0.148	0.133	0.136	0.109	0.099	0.092

Notes: Observations are adult respondents. The dependent variable in column 1 is a dummy =1 if the respondent says they would cut their daughter if 100 out of 100 in the community decided to cut their daughters; in column 2 is a dummy =1 if the respondent says they would cut their daughter if 80 out of 100 did so; in column 3 is a dummy =1 if the respondent says they would cut their daughter if 60 out of 100 did so; in column 4 is a dummy =1 if the respondent says they would cut their daughter if 40 out of 100 did so; in column 5 is a dummy =1 if the respondent says they would cut their daughter if 20 out of 100 did so; in column 6 is a dummy =1 if the respondent says they would cut their daughter if no one out of 100 did so. All regressions control for district fixed effects, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level and occupation. Standard errors are clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

B List experiment

For the list experiment, we ask the respondent to report the number of items (from a list of 4 or 5 items) that applies to them. In particular, respondents are read the following script: "Now I'm going to read some statements about many different things. Some of these statements will be true and some will not. After I read all statements, please tell me how many of them are true for you. I don't want to know which ones, just how many." In order to make sure the respondents understand the exercise, we first conduct a practice round. They are told "Let's try with an example first. Suppose I read you 4 statements. After I read each statement, I want you to count with your fingers if it is true, and keep track without showing to me. I will then ask you how many are true. Let me demonstrate first. I will now read you the statements. As I'm reading, count with your fingers without showing me. At the end, you'll tell me how many are true for you." Then enumerators read 4 statements and respondents state how many of these statements are true for them. Following this exercise round, the respondents are provided with the actual statements.

Every respondent is randomly assigned to a group (group "A" or "B"). Tables B.1, B.2, and B.3 display the balance checks between both groups. Respondents in group A are presented a list of 5 items, one of which is the sensitive item.; while respondents in group B are presented a list of 4 items, without the sensitive item. Respondents in group A are asked to state how many of the following statements are true: "1) You like to watch TV, 2) You have been to the market yesterday, 3) You have a daughter who has been Pharaonic cut, 4) You have never been to Mogadishu, 5) You have two brothers"; while respondents in group A are only asked to state how many of the 4 non-sensitive statements are true (Table B.4). The only difference between the two lists is the presence of the statement "You have a daughter who has been Pharaonic cut" in the list presented to group A but not in group B. In order to calculate the percentage of people for whom the sensitive item is true (i.e., the percentage of people who have a daughter who has been Pharaonic cut), one can look at the difference between the average number of statements reported as "true" by respondents in groups A relative to B. Since the assignment of individuals to group A or B is random, this implies that there is no reason why the number of true statements in the two groups should be different, other than the presence of the sensitive item. Therefore, the difference in the average number of "true" statements for respondents in group A vs. B is an estimate of the share of respondents for whom the sensitive statement is "true".

Table B.5 shows individual-level regressions for the list experiment. The dependent variable is the number of true statements of each respondent. The variable "sensitive item" is a dummy that takes value 1 if the respondent was assigned to the group with the sensitive item (group A). Its coefficient can be interpreted as the 'true' share of respond-

ents belonging to the control group that have a Pharaonic-cut daughter. The coefficients on the interaction terms between the sensitive item and the treatment dummies represent the treatment effects on the 'true' share of respondents that have a Pharaonic-cut daughter. (Note that the coefficients of the standalone treatment dummies do *not* have a meaningful interpretation: they are the treatment effects on the number of true non-sensitive items, which –as expected– are null.

Table B.1: Balance Table of List Experiments (A): Baseline characteristics and demographic characteristics

	(1)		(2)		(3)	(4)
	Group A		**Group B**		**P-value**	**Norm diff.**
	N	Mean/SD	N	Mean/SD		
Balistbine guesses characteristics						
Guess of sh. of own group that supp. abandon Phar.	1924	0.736	2038	0.748	0.066*	0.037
		(0.242)		(0.237)		
Guess of sh. of other group that supp. abandon Phar.	1924	0.654	2038	0.661	0.411	0.019
		(0.263)		(0.260)		
Guess of sh. of grandmothers that supp. abandon Phar.	1924	0.595	2038	0.583	0.253	-0.028
		(0.294)		(0.293)		
Guess of sh. of grandfathers that supp. abandon Phar.	1924	0.603	2038	0.608	0.564	0.013
		(0.288)		(0.286)		
Demographic characteristics						
Female	1924	0.495	2038	0.521	0.100	0.037
		(0.500)		(0.500)		
Age	1924	37.771	2038	38.083	0.449	0.016
		(13.783)		(14.413)		
Number of sons	1924	2.943	2038	2.833	0.104	-0.034
		(2.330)		(2.237)		
Number of daughters	1924	2.880	2038	2.773	0.110	-0.034
		(2.300)		(2.233)		
Ever married	1913	0.923	2027	0.917	0.513	-0.014
		(0.267)		(0.276)		
Age at marriage	1759	21.249	1852	21.137	0.618	-0.012
		(6.681)		(6.674)		
Married a relative	1760	0.395	1853	0.371	0.166	-0.035
		(0.489)		(0.483)		
Woman brought any asset into marriage	908	0.083	1008	0.070	0.272	-0.032
		(0.275)		(0.256)		
Subclan: Majarteen (Darood)	1924	0.270	2038	0.291	0.111	0.033
		(0.444)		(0.455)		
Subclan: Others (Darood)	1924	0.177	2038	0.166	0.386	-0.019
		(0.382)		(0.372)		
Subclan: Garxajis (Dir)	1924	0.173	2038	0.157	0.169	-0.031
		(0.378)		(0.363)		
Subclan: Habar Awal (Dir)	1924	0.133	2038	0.139	0.542	0.013
		(0.340)		(0.346)		
Subclan: Habar Jeclo (Dir)	1924	0.127	2038	0.124	0.765	-0.007
		(0.333)		(0.329)		
Subclan: Other (Dir)	1924	0.069	2038	0.068	0.853	-0.004
		(0.254)		(0.251)		
Subclan: Others	1924	0.051	2038	0.054	0.699	0.010
		(0.221)		(0.227)		

Table B.2: Balance Table of List Experiments (B): Socioeconomic characteristics

	(1)		(2)		(3)	(4)
	Group A		**Group B**		**P-value**	**Norm diff.**
	N	Mean/SD	N	Mean/SD		
Socioeconomic characteristics						
Asset index	1924	0.158	2038	0.152	0.885	-0.003
		(1.238)		(1.226)		
HH runs non-agricultural business	1924	0.125	2038	0.114	0.299	-0.024
		(0.331)		(0.318)		
Education: No formal education	1924	0.457	2038	0.469	0.424	0.017
		(0.498)		(0.499)		
Education: Incomplete primary	1924	0.199	2038	0.186	0.284	-0.022
		(0.399)		(0.390)		
Education: Primary, incomplete secondary	1924	0.093	2038	0.086	0.399	-0.018
		(0.291)		(0.280)		
Education: Secondary or above	1924	0.100	2038	0.108	0.345	0.019
		(0.300)		(0.311)		
Education: Koranic study	1924	0.121	2038	0.122	0.949	0.001
		(0.326)		(0.327)		
Occupation: day-laborer	1924	0.305	2038	0.318	0.313	0.020
		(0.460)		(0.466)		
Occupation: farmer	1924	0.128	2038	0.118	0.396	-0.022
		(0.334)		(0.322)		
Occupation: self-employed (non-agricultural)	1924	0.072	2038	0.059	0.156	-0.035
		(0.258)		(0.236)		
Occupation: salaried work	1924	0.076	2038	0.070	0.461	-0.016
		(0.265)		(0.255)		
Occupation: housework	1924	0.174	2038	0.180	0.632	0.010
		(0.379)		(0.384)		
Occupation: student	1924	0.034	2038	0.035	0.845	0.004
		(0.181)		(0.183)		
Occupation: other	1924	0.207	2038	0.216	0.483	0.016
		(0.406)		(0.412)		

Table B.3: Balance Table of List Experiments (C): Household head characteristics and district of origin

	(1)		(2)		(3)	(4)
	Group A		**Group B**		**P-value**	**Norm diff.**
	N	Mean/SD	N	Mean/SD		
Household head characteristics						
Age	1917	42.553	2033	42.696	0.735	0.007
		(14.037)		(14.466)		
Education: No formal education	1924	0.455	2038	0.467	0.448	0.017
		(0.498)		(0.499)		
Education: Incomplete primary	1924	0.192	2038	0.180	0.297	-0.021
		(0.394)		(0.384)		
Education: Primary, incomplete secondary	1924	0.092	2038	0.086	0.482	-0.015
		(0.289)		(0.280)		
Education: Secondary or above	1924	0.104	2038	0.117	0.143	0.028
		(0.306)		(0.321)		
Education: Koranic study	1924	0.126	2038	0.122	0.615	-0.010
		(0.332)		(0.327)		
Occupation: day-laborer	1924	0.357	2038	0.383	0.072*	0.038
		(0.479)		(0.486)		
Occupation: farmer	1924	0.132	2038	0.118	0.260	-0.029
		(0.339)		(0.323)		
Occupation: self-employed (non-agricultural)	1924	0.067	2038	0.054	0.160	-0.037
		(0.250)		(0.227)		
Occupation: salaried work	1924	0.089	2038	0.079	0.205	-0.024
		(0.285)		(0.271)		
Occupation: housework	1924	0.134	2038	0.141	0.525	0.014
		(0.341)		(0.348)		
Occupation: student	1924	0.021	2038	0.024	0.431	0.016
		(0.143)		(0.153)		
Occupation: other	1924	0.197	2038	0.198	0.925	0.002
		(0.398)		(0.399)		
District of origin						
Hargeisa	1924	0.207	2038	0.213	0.632	0.010
		(0.406)		(0.409)		
Burao	1924	0.170	2038	0.167	0.850	-0.005
		(0.376)		(0.373)		
Erigavo	1924	0.155	2038	0.138	0.149	-0.034
		(0.362)		(0.345)		
Badhan	1924	0.149	2038	0.135	0.182	-0.028
		(0.356)		(0.342)		
Southern Galkayo	1924	0.168	2038	0.175	0.572	0.012
		(0.374)		(0.380)		
Galdogob	1924	0.150	2038	0.171	0.043**	0.042
		(0.357)		(0.377)		

Table B.4: Item Lists

List 1	
Group A	**Group B**
1. You like to watch TV	1. You like to watch TV
2. You have been to the market yesterday	2. You have been to the market yesterday
3. You have a daughter who has been Pharaonic cut	3. You have two brothers
4. You have two brothers	4. You have never been to Mogadishu
5. You have never been to Mogadishu	

Table B.5: List Experiment and Stated Measure

	(1) Has a Pharaonic cut daughter (inferred, list)
T1 : Poll + Announcement	0.005
	(0.085)
T2 : Poll + Coordination	-0.081
	(0.084)
T3 : Poll + Announce. + Coord.	0.089
	(0.081)
Sensitive item	0.324***
	(0.069)
T1 × Sensitive item	-0.184*
	(0.096)
T2 × Sensitive item	0.002
	(0.107)
T3 × Sensitive item	-0.153
	(0.094)
Mean in control	2.117
Observations	3,962
R^2	0.090

Notes: Observations are adult respondents. The dependent variable in column 1 is the number of statements in the list that respondents say that are correct (ranging from 0-4 in the group without the sensitive item and 0-5 in the group with the sensitive item). A positive coefficient in the group with the sensitive item measures the share of respondents for which the sensitive item is true (i.e., "I have a daughter Pharaonic cut"). If the whole sample disagrees with the sensitive item, its coefficient should be equal to zero. The interaction between the treatment dummy and the sensitive item reports a truthful measure of the treatment effects on the relevant outcomes. All regressions control for district fixed effects, respondent's age, respondent's sex, number of sons, number of daughters, respondent's education level, respondent's marital status, respondent's occupation, respondent's subclan, household asset index, a dummy =1 if anyone in the household runs a non-agricultural enterprise, a dummy =1 if the respondent is the head of the household, household head's age, education level, and occupation categories. Robust standard errors are reported (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).