

Changing Cultural Attitudes on Female Genital Cutting

Sonja Vogt^{1,*}, Nadia Ahmed Mohammed Zaid²,

Hilal El Fadil Ahmed³, Ernst Fehr^{1,†}, & Charles Efferson^{1,*†}

¹*Department of Economics, University of Zurich*

²*Omdurman, Sudan*

³*Khartoum, Sudan*

**These authors contributed equally to this research.*

†Shared senior authorship.

As globalisation brings people with incompatible attitudes into contact, cultural conflicts inevitably arise. Little is known about how to mitigate conflict and about how the conflicts that occur shape the cultural evolution of the groups involved. Female genital cutting provides a prominent example¹⁻³. Governments and international agencies have promoted the abandonment of cutting for decades, but the practice remains widespread with attendant health risks for millions of girls and women^{4,5}. In their efforts to end cutting, international actors have often adopted the view that cutting is locally pervasive and entrenched¹. This implies the need to introduce values and expectations from outside the local culture. Members of the target society may view such interventions as unwelcome intrusions^{1-3,6-9}, and campaigns promoting abandonment have sometimes led to backlash^{1,7,8,10,11} as they struggle to reconcile cultural tolerance with the conviction that cutting violates universal human rights^{1,9}. Cutting, however, is not necessarily locally pervasive and entrenched^{1,3,12}. We designed ex-

periments on cultural change that exploit the existence of conflicting attitudes within cutting societies. We produced four entertaining movies that served as treatments in two experiments in Sudan, and we developed an implicit association test to unobtrusively measure attitudes about cutting. The movies depart from the view that cutting is locally pervasive by dramatising members of an extended family as they confront each other with divergent views about whether the family should continue cutting. The movies significantly improved attitudes about uncut girls, with one in particular having a relatively persistent effect. These results show that using entertainment to dramatise locally discordant views can provide a basis for applied cultural evolution without accentuating inter-cultural divisions.

Ethnographic work emphasises that cutting is subject to constant negotiation and renegotiation within the culture in question^{2,3,7,8}. If true, the seeds of abandonment should already be present within a cutting community. Recent analyses of large data sets have confirmed this view by showing that attitudes associated with cutting vary tremendously at the household and individual levels¹²⁻¹⁴. Significantly, individual variation in attitudes within households^{3,13} is effectively as local as possible, and programmes that take this kind of local heterogeneity as a point of departure could provide a basis for change³. Put differently, the international movement to reduce violence against women and girls could be effective to the extent that it is translated into local terms². Locally heterogeneous views about cutting offer an opportunity to do this by casting the debate about cutting versus abandonment into a local vernacular^{2,3}.

To see if this would work, we produced four telenovela-style movies that portray locally dis-

parate views on cutting. As explained below, the movies vary systematically in terms of content, with some movies examining the link between cutting and individual values, and some movies examining the effects of cutting on the marriage prospects of young women¹⁵. The movies were shown to people in two fully randomised and controlled experiments. The variation in content across movies allowed us to see if certain arguments related to cutting are more compelling than others by systematically decomposing and manipulating the content of interest across several treatments.

In addition, our task was to see if entertainment focused on local heterogeneity can change attitudes about cutting. The history of entertainment in relation to gender bias dates back at least to a Peruvian telenovela in 1969¹⁶, and this history includes a handful of documentaries and radio dramas that have addressed cutting^{17,18}. Only recently, however, have social scientists developed methods to identify causal effects, as opposed to mere correlations, associated with entertainment. Existing causal studies based on entertainment have not addressed female genital cutting¹⁹⁻²², and more broadly our study is one of very few to allow causal inferences about programmes of any kind related to cutting²³.

In accord with the focus on local heterogeneity, the movies in our experiments take a non-judgemental approach to cutting. They do not consistently present arguments in favour of abandonment, and they do not associate negative characters with the support of cutting. Rather, the movies dramatise how difficult a decision about cutting can be for parents who want the best for their daughters in a society where cutting is common, but attitudes and practices vary. Our focus on

local heterogeneity also works around a challenge that typically arises when using entertainment to change attitudes. Attitudinal change via entertainment faces a basic trade-off. It must present people with characters and situations that are identifiably familiar but still provide examples of new ways to think¹⁶. Research has not revealed much about how to negotiate this trade-off¹⁶. We work around the trade-off by allowing characters within an extended family to disagree about cutting. Importantly, the members of the family do not choose cutting or abandonment independently of each other. They do, however, vigorously debate the issue. This reflects the individual variation in attitudes known to exist¹³ and situates the cultural tension between cutting and abandonment as locally as possible^{2,3}.

We produced four movies that served as experimental treatments (Supplementary Information, § 2). The movies differ in terms of the arguments the fictional characters articulate when expressing their views on cutting (Methods). A number of hypothesised motives for cutting have been proposed, but broadly speaking they fall into two categories¹⁵. First, people may have opinions for or against cutting because people hold individual values related to health, purity, and perceived religious obligations. For opinions of this sort, cutting has an intrinsic value, positive or negative, rooted in the individual's view of the practice. Second, people may have opinions for or against cutting because they hold certain beliefs about how cutting will affect the future marriage prospects of their daughters. In this sense, cutting has an extrinsic value widely assumed to depend on whether families expect cut wives for their sons. We varied whether the characters in the movies expressed opinions in these two categories. We did so in all combinations, which gave us a **values** movie, a **marriageability** movie, a **combined** movie that addresses both values and

marriageability, and a **control** movie.

All movies are about the same extended family in a rural area in Sudan. The extended family includes multiple couples and their children, a venerable grandfather, and other relatives and spouses. Family members live together in a large compound. The movies share a main plot that is not about cutting. The main plot involves a heady mix of love, intrigue, deception, forgiveness, and engaging personalities. The main plot is always the same, and it consists of nearly identical scenes across all four movies. The control includes only the main plot. The three remaining movies include subplots related to female genital cutting. For the subplots, several couples in the extended family have daughters approaching the age of cutting, and family members debate whether they should cut their daughters or abandon the practice. The three subplots consist of parallel scenes and dialogues. A given scene includes a set of characters having a discussion or dispute about cutting. Regardless of treatment, the same characters are always in a scene, and the scene always has the same position in the narrative structure of movie. Only the content of dialogues varies by treatment. All movies are approximately 90 minutes long. Subplots about cutting constitute approximately 30% of this (c. 27 minutes).

The movies do not simply display conflict and leave disputes unresolved. All three movies about cutting provide examples of how to negotiate conflicting views within a family. The movies reveal that husbands and wives discuss cutting together. This is an important innovation in Sudan where people can have difficulty discussing cutting openly when men and women are both present. Moreover, the three movies about cutting end with the family presenting the idea of abandoning

the practice to the ageing grandfather, the senior member of the extended family. This scene is the culmination of one key source of conflict in the movies. Specifically, conflicts occur repeatedly because some feel that family members must approach the grandfather to discuss cutting, while others feel that doing so will cause him unnecessary stress. When the family finally does approach the grandfather, he approves of the proposal to abandon cutting. He speculates that his dead wife would also approve, and he draws an analogy between cutting and facial scarring, a practice once common in Sudan but now in rapid decline. The grandfather's reaction seals the family's decision to stop cutting, and in this way the movies provide a model for how to include young and old generations in debates about cutting. In keeping with our general strategy, the final scenes with the grandfather are parallel across the three movies about cutting; only the details of the discussion vary.

We used the four movies as treatments in two experiments. For the first experiment (Supplementary Information, § 3), we randomly assigned individuals within communities to the four treatments, and we collected attitudinal data using an implicit association test (Supplementary Information, § 1) immediately after participants watched the movies. By minimising socially desirable responses, implicit association tests have been used to successfully measure attitudes about many sensitive topics in which respondents may not be prepared to explicitly reveal their attitudes²⁴. For the second experiment (Supplementary Information, § 4), we combined 122 communities into 88 groups of communities by assigning communities within three kilometres of each other to the same group. We randomly assigned these 88 groups to the four treatments. In each community, we collected baseline attitudinal data using the implicit association test, and several weeks later we

screened the appropriate movie one time. We collected follow-up attitudinal data approximately one week (5-24 days, $\bar{x} = 6.46$, with 6 and 7 days accounting for more than 90% of observations) after the movie screening.

The two experiments complement each other methodologically. By randomising at the individual level, we conducted the first experiment in a given community in about three hours. This maximised experimental control. All participants watched their assigned movies, and we were thus able to treat everyone we intended to treat. Moreover, apart from one participant who did not complete the implicit association test, we have data for everyone who watched a movie. Most importantly, we were able to see if the movies had an immediate effect on attitudes.

Randomising within communities, however, is not well suited to identifying effects over a longer time horizon. Had we returned to measure attitudes after one week, for example, participants assigned to different treatments within communities would have presumably discussed the movies they saw, and the treatments would have contaminated each other as a result. The second experiment solved this problem by assigning entire groups of communities in a given geographical area to a single treatment.

However, because the second experiment involved three activities over several weeks, not all recruited participants saw a movie. Movie attendance was very high but not perfect (80.2-87.1%, Supplementary Information, Table S4). This means we must distinguish between effects associated with intending to treat participants and effects associated with participants who actually saw their assigned movies (Supplementary Information, § 4). Furthermore, participation in the two data

collection events was high, but again not perfect (88.4-93.9%, Supplementary Information, Table S4). Consequently, we have missing observations, and we use both inverse probability weighting²⁵ and multiple imputation²⁶ to ensure that results are robust to missing observations (Supplementary Information, § 4).

Our implicit association test measures attitudes about cut versus uncut girls (Supplementary Information, § 1). The test produces D scores for individual participants, where $D \in [-2, 2]$. A D score in $[-2, 0)$ indicates that the subject had relatively negative associations with uncut girls, while a score in $(0, 2]$ means the subject had relatively positive associations with uncut girls. $D = 0$ indicates that the subject had no relative implicit associations. Implicit association tests have high predictive validity when used, as in our experiments, to address socially sensitive topics with subjects who have little prior experience with such measures²⁷. Implicit attitudinal measures have also been shown to reduce or eliminate social desirability bias when measuring attitudes related to female genital cutting²⁸. We validated our test by showing (Extended Data Fig. 1) that average implicit attitudes by community were highly significantly correlated with cutting rates by community across 45 communities in Gezira (Pearson's correlation, $\rho = -0.423$, one-sided $p = 0.0008$, based on a two-dimensional weighted bootstrap explained in the Supplementary Information, § 1).

After the implicit association test, each participant completed a short questionnaire in which we collected data on gender, age, endogamy, whether the participant had sons or daughters, and whether the participant and her family had a history of nomadism. In the area in Gezira where we worked, two common forms of production are farming and herding. One hypothesised motive

for cutting and infibulation is to reduce fears about paternity uncertainty⁶. If this motive is important, we predicted that high mobility would have made such fears more important among herders than among farmers, and families with a history of nomadism would have had relatively negative attitudes about uncut girls. Apart from age and sex, all data were collected under strictly anonymous laboratory conditions (Supplementary Information, § 1), which can be extremely effective at reducing social desirability biases when researching culturally sensitive issues²⁹.

For the first experiment, none of the control variables were significantly related to attitudes (Extended Data Table 1). The three movies about cutting all produced large, robust, and highly significant increases in positive attitudes about uncut girls. The three experimental treatments improved attitudes towards uncut girls with effects of 55-64% of a standard deviation (Extended Data Table 1, $p \leq 0.001$ or $p \in (0.001, 0.01]$). Only 35.6% of participants in the control had positive associations with uncut girls ($D > 0$), while 66.7-72.9% of participants had associations of this sort in the three experimental treatments. The effects of the three experimental treatments were indistinguishable (Fig. 1a and Extended Data Table 1). Arguments related to individual values, marriageability, and a combination of the two all led to similar improvements in attitudes towards uncut girls. This finding suggests that individual values and marriageability concerns both influence attitudes about cutting and attitudinal change.

Results from the second experiment confirm the joint importance of individual values and marriageability. We found a robust and highly significant improvement in attitudes about uncut girls, but only for the combined movie. The combined movie improved implicit association scores

with effects of approximately 10-11% of a standard deviation (Extended Data Table 2 with all $p \in (0.001, 0.01]$ and Extended Data Table 3 with $p \in (0.001, 0.01]$ or $p \in (0.01, 0.05]$). Interestingly, the effects of the combined movie were heterogeneous (Extended Data Tables 4-7). In a follow-up analysis, we ranked participants according to their baseline implicit association scores and divided participants at the median. The effect of the combined treatment was largely due to changes among those participants with baseline scores above the median, namely those participants less negative about uncut girls when the experiment began (Extended Data Tables 6-7, all $p \in (0.01, 0.05]$).

These results hold when estimating both the intention to treat (Extended Data Tables 2 and 6), which does not account for whether a participant actually saw her assigned movie, and the effects of the treatments on the movie-goers (Extended Data Tables 3 and 7). Moreover, regardless of how we handle missing observations, effect sizes and significance levels are extremely stable (Extended Data Tables 2-7).

Apart from the effect associated with the combined treatment, we also found (Extended Data Table 2) highly significant positive effects associated with being a female (all $p \leq 0.001$) and with age (all $p \in (0.001, 0.01]$), as well as significant positive effects associated with having daughters (all $p \in (0.01, 0.05]$). This means that, all else equal, women were more positive about uncut girls than men, older participants were more positive than younger participants, and participants with daughters were more positive than those without. We additionally found (Extended Data Table 2) large and highly significant negative effects associated with nomadism. The negative effect associated with participants identifying themselves as nomads was about 21% of a standard

deviation (all $p \leq 0.001$), and the negative effect associated with having ancestors who were nomads was about 15% of a standard deviation (all $p \leq 0.001$). This result suggests that a history of paternity concerns⁶ could be contributing to support for cutting among participants with a nomadic lifestyle.

Our results show that movies emphasising local heterogeneity in terms of both values and marriageability can change attitudes to favour the abandonment of cutting. Three important implications follow. First, programmes that take local heterogeneity as a point of departure^{2,3} offer a promising avenue for cultural change when a tension exists between human rights, as an international directive, and a local culture. Our experiments show that dramatising discordant views on cutting within a family can improve attitudes about uncut girls. Such an approach could be applicable beyond Gezira, Sudan, because empirical evidence reveals that local heterogeneity in attitudes related to cutting is widespread^{3,12,13}. However, to the extent that some areas are characterised by locally homogeneous support for cutting (Supplementary Information, § 5), a focus on local heterogeneity might be less effective than what we observed.

Second, addressing individual values and marriageability concerns together produces attitudinal changes that last longer than those associated with addressing values and marriageability separately. The link between cutting and marriageability has been an influential hypothesis⁶, and our results confirm that marriageability can be critical. However, our results indicate that exposing variation in individual values is also critical for changing cultural attitudes.

Finally, our results suggest that entertainment, by changing attitudes, could contribute to the

abandonment of cutting. The demand for entertainment is enormous and seemingly ubiquitous²². Messaging related to cutting can thus be incorporated, for example, in television shows that are conceived from the beginning as for-profit ventures. To support the feasibility of this idea, recent studies reveal that commercial television has improved the status of women in India¹⁹ and Brazil²⁰. Consequently, even if movies and television are expensive to produce and distribute, the demand for entertainment might defray much of the cost. Furthermore, because the demand for entertainment is so widespread, messaging embedded in popular entertainment could potentially reach a broad cross section of the population. This stands in contrast, for example, to non-fiction documentaries, which mainly reach those already amenable to the message²².

Importantly, the distinguishing characteristics of entertainment could offset the limitations our results reveal in terms of attitudinal change. Our first experiment showed large immediate improvements in attitudes about uncut girls after approximately 27 minutes of material related to cutting. Our second experiment, however, showed smaller effects after one week. Sustained exposure to appropriate messaging could thus be essential for long-term change. If so, covering costs via the demand for entertainment is especially attractive. Our second experiment also revealed that participants who entered the experiment with the most negative attitudes about uncut girls did not change their attitudes. Again, sustained exposure might lead to change among resistant segments of population, but one must reach these segments in the first place. The appeal of entertainment is again attractive because it limits the risk that the messaging never reaches those most in favour of cutting²². By positioning the debate about cutting locally, we found that entertaining movies addressing both individual values and marriageability can lead people to have more positive attitudes

about uncut girls and provide a means of gently provoking change from within a cutting culture.

1. Shell-Duncan, B. & Hernlund, Y. Female “Circumcision” in Africa: Dimensions of the Practice and Debates. In Shell-Duncan, B. & Hernlund, Y. (eds.) *Female “Circumcision” in Africa: Culture, Controversy, and Change*, 1–40 (Boulder, CO: Lynne Rienner, 2000).
2. Merry, S. E. *Human Rights and Gender Violence: Translating International Law Into Local Justice* (Chicago: University of Chicago Press, 2006).
3. Hernlund, Y. & Shell-Duncan, B. Contingency, context, and change: Negotiating female genital cutting in the gambia and senegal. *Africa Today* **53**, 43–57 (2007).
4. Bishai, D. *et al.* Estimating the obstetric costs of female genital mutilation in six African countries. *Bulletin of the World Health Organization* **88**, 281–388 (2010).
5. World Health Organization. Female genital mutilation. http://www.who.int/topics/female_genital_mutilation/en/ (2014, accessed 16 September 2015).
6. Mackie, G. Female genital cutting: the beginning of the end. In Shell-Duncan, B. & Hernlund, Y. (eds.) *Female “Circumcision” in Africa: Culture, Controversy, and Change*, 253–283 (Boulder, CO: Lynne Rienner Publishers, 2000).
7. Gruenbaum, E. *The Female Circumcision Controversy: An Anthropological Perspective* (Philadelphia: University of Pennsylvania Press, 2001).

8. Boddy, J. *Civilizing Women: British Crusades in Colonial Sudan* (Princeton: Princeton University Press, 2007).
9. Shell-Duncan, B. From health to human rights: Female genital cutting and the politics of intervention. *American Anthropologist* **110**, 225–236 (2008). URL <http://dx.doi.org/10.1111/j.1548-1433.2008.00028.x>.
10. Thomas, L. “Ngaitana (I Will Circumcise Myself)”: Lessons from Colonial Campaigns to Ban Excision in Meru, Kenya. In Shell-Duncan, B. & Hernlund, Y. (eds.) *Female “Circumcision” in Africa: Culture, Controversy, and Change*, 129–150 (Boulder, CO: Lynne Rienner, 2000).
11. Camilotti, G. Interventions to stop female genital cutting and the evolution of the custom: Evidence on age at cutting in Senegal. *Journal of African Economies* (2015). URL <http://jae.oxfordjournals.org/content/early/2015/07/20/jae.ejv013.abstract>. <http://jae.oxfordjournals.org/content/early/2015/07/20/jae.ejv013.full.pdf+html>.
12. Shell-Duncan, B. & Hernlund, Y. Are there “stages of change” in the practice of female genital cutting? qualitative research findings from Senegal and the Gambia. *African Journal of Reproductive Health* **10**, 57–71 (2006).
13. Bellemare, M. F., Novak, L. & Steinmetz, T. L. All in the family: Explaining the persistence of female genital cutting in West Africa. *Journal of Development Economics* **116**, 252 – 265 (2015). URL <http://www.sciencedirect.com/science/article/pii/S0304387815000620>.

14. Efferson, C., Vogt, S., Elhadi, A., Ahmed, H. E. F. & Fehr, E. Female genital cutting is not a social coordination norm. *Science* **349**, 1446–1447 (2015).
15. Shell-Duncan, B., Wander, K., Hernlund, Y. & Moreau, A. Dynamics of change in the practice of female genital cutting in Senegambia. *Social Science & Medicine* **73**, 1275–1283 (2011).
16. La Ferrara, E. Mass media and social change: Can we use television to fight poverty? *CEPR Discussion Paper No. DP10954* (2015). URL <http://ssrn.com/abstract=2697567>.
17. Greiner, K., Singhal, A. & Hurlburt, S. With an antenna we can stop the practice of female genital cutting: a participatory assessment of *Ashreat Al Amal*, an entertainment-education radio soap opera in Sudan. *Investigación y Desarrollo* **15**, 226–259 (2007).
18. Love Productions. The Cruel Cut. <http://leylahussein.com/the-cruel-cut/> (2014, accessed 25 May 2016).
19. Jensen, R. & Oster, E. The power of TV: Cable television and women’s status in India. *The Quarterly Journal of Economics* **124**, 1057–1094 (2009).
20. La Ferrara, E., Chong, A. & Duryea, S. Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics* **4**, 1–31 (2012). URL <http://www.aeaweb.org/articles.php?doi=10.1257/app.4.4.1>.
21. Kearney, M. S. & Levine, P. B. Media influences on social outcomes: The impact of MTV’s 16 and pregnant on teen childbearing. *American Economic Review* **105**, 3597–3632 (2015). URL <http://www.aeaweb.org/articles.php?doi=10.1257/aer.20140012>.

22. DellaVigna, S. & La Ferrara, E. Economic and social impacts of the media. In Anderson, S., Waldfogel, J. & Stromberg, D. (eds.) *Handbook of Media Economics* (Amsterdam: Elsevier, Forthcoming).
23. Berg, R. C. & Denison, E. Effectiveness of interventions designed to prevent female genital mutilation/cutting: A systematic review. *Studies in Family Planning* **43**, 135–146 (2012). URL <http://dx.doi.org/10.1111/j.1728-4465.2012.00311.x>.
24. Nosek, B. A., Greenwald, A. G. & Banaji, M. R. The implicit association test at age 7: a methodological and conceptual review. In Bargh, J. A. (ed.) *Automatic Processes in Social Thinking and Behavior*, 265–292 (Psychology Press, 2007).
25. Robins, J. M., Rotnitzky, A. & Zhao, L. P. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association* **89**, 846–866 (1994).
26. Schafer, J. L. *Analysis of Incomplete Multivariate Data* (Boca Raton, FL: Chapman & Hall, 1997).
27. Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L. & Banaji, M. R. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology* **97**, 17–41 (2009).
28. de Cao, E. & Lutz, C. Measuring attitudes regarding female genital mutilation through a list experiment. Tech. Rep. (2015).

29. Chauchard, S. Using mp3 players in surveys: The impact of a low-tech self-administration mode on reporting of sensitive attitudes. *Public Opinion Quarterly* **77**, 220–231 (2013).

Supplementary Information Available with the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Gezira State government, the Sudan Nation Council for Child Welfare, the Gezira State Council for Child Welfare, and local authorities in Umalgoura and East Gezira for supporting the study. We would further like to acknowledge the efforts of the Gezira State Council for Child Welfare, our many facilitators and participants, and the national field office of UNICEF in Khartoum. We thank Omer Abdazeez Abdarazig Abdalla, Waleed Omer Babikir Alalfi, and Abduelrahman Al Shibli for their contributions to the materials for the IAT, and we thank Walid Omer Babikir Alalfi for his screenplays. Additional thanks go to Mahzarin Banaji, Pietro Biroli, and Rainer Winkelmann for valuable feedback. This study was funded primarily by the Swiss National Committee of UNICEF. It is also part of a European Research Council grant on the Foundations of Economic Preferences. Funding agencies played no role in the design of the study, data collection, data analysis and interpretation, or the writing of the paper.

Author Contributions SV, EF, and CE initiated the study. SV and CE designed the experiments with input from EF. SV and CE contributed to the writing and production of the movies and developed the implicit association test for measuring attitudes. NAMZ pre-tested the movies. SV, NAMZ, HEFA, and CE recruited participants and liaised with government and community officials. SV, NAMZ, HEFA, and CE planned and conducted the experiments. CE analysed the data with input from SV and EF. SV, EF, and CE interpreted the results and wrote the paper.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence should be addressed to SV (sonja.vogt@econ.uzh.ch), EF (ernst.fehr@econ.uzh.ch), and CE (charles.efferson@econ.uzh.ch).

Figure 1 | Treatment effects for unnormalised implicit association scores. Larger scores indicate relatively positive attitudes about uncut girls. **a**, Experiment 1, mean scores as measured immediately after movies (Extended Data Table 1). Error bars show 95% bootstrapped confidence intervals clustered on 40 movie groups, where each group consisted of three to five subjects who watched a movie together. Clustering accounts for dependence due to social interactions during the movies. Data are from 189 participants in five communities. **b**, Experiment 2, mean scores from the follow-up data collection, approximately one week after the movies, minus mean scores from the baseline data collection. Error bars show 95% bootstrapped confidence intervals clustered on 88 groups of communities assigned to treatments. Data are from 7729 participants in 122 communities. Difference-in-difference estimations (Supplementary Information, § 4) consistently produce significant effects for the combined treatment of 10.1-11.3% of a standard deviation (Extended Data Tables 2-3 and 6-7).

Extended Data Figure 1 | Correlation between attitudes and cutting practices at the community level. For 45 communities in the state of Gezira, Sudan, the figure shows the relationship between estimated cutting rates and average scores by community¹⁴ from the implicit association test (Pearson's correlation, $\rho = -0.423$, one-sided $p = 0.0008$, based on a two-dimensional weighted bootstrap explained in the Supplementary Information, § 1). The unweighted least squares line is shown for reference.

Extended Data Table 1 | Experiment 1, regression results for the implicit association test. The response variable consists of implicit association scores normalised to have a mean of zero and a standard deviation of one. Standard errors were calculated by clustering on movie group, a procedure that accounts for any dependence due to social interactions while watching the movies. Treatment coefficients show large, significant improvements in participant attitudes towards uncut girls. The three experimental treatments are indistinguishable from each other (Model without controls: Values = Marriageability, $F = 0.169$, $p = 0.682$; Values = Combined, $F = 0.001$, $p = 0.973$; Marriageability = Combined, $F = 0.212$, $p = 0.646$). Because two participants did not complete the anonymous questionnaire after the implicit association test, the model with controls is fit to 187 observations, while the model without controls is fit to all 189 observations.

Extended Data Table 2 | Experiment 2, intention-to-treat effects estimated by difference-in-difference. The response variable consists of implicit association scores normalised to have a mean of zero and a standard deviation of one. Interactions between the follow-up data collection and the treatment dummies estimate the difference-in-difference effects relative to the control (e.g. Follow-up \times Combined; Supplementary Information, § 4). These effects are highlighted in blue. To ensure results robust to missing data, we analysed the unweighted complete cases, the inverse probability weighting of complete cases, and multiple imputations. The complete cases include 13906 observations from 7729 participants. Under multiple imputation, we imputed 1552 observations for each of 200 imputations. For treatment assignment, 122 communities were combined into 88 groups according to geographic proximity. Robust standard errors were calculated by clustering on these 88 groups.

Extended Data Table 3 | Experiment 2, the effects of the treatments on the movie-goers. The response variable consists of implicit association scores normalised to have a mean of zero and a standard deviation of one. Estimates show the difference-in-difference effects by treatment relative to the control (Supplementary Information, § 4). To check for robustness, we analysed the unweighted complete cases, the inverse probability weighting of complete cases, and multiple imputations. Data include 13906 observations from 7729 participants. Under multiple imputation, we imputed 1552 observations for each of 200 imputations. For treatment assignment, 122 communities were combined into 88 groups according to geographic proximity. Robust standard errors were calculated by clustering on these 88 groups.

Extended Data Table 4 | Experiment 2, intention-to-treat effects for subjects with a baseline implicit association score at or below the median. The response variable consists of implicit association scores normalised to have a mean of zero and a standard deviation of one. Interactions between the follow-up data collection and the treatment dummies (e.g. Follow-up \times combined; Supplementary Information, § 4) estimate the difference-in-difference effects relative to the control. These effects are highlighted in blue. To ensure results robust to missing data, we analysed the unweighted complete cases, the inverse probability weighting of complete cases, and multiple imputations. Data include 6659 observations from 3541 participants. Under multiple imputation, we imputed 423 observations for each of 200 imputations. For treatment assignment, 122 communities were combined into 88 groups according to geographic proximity. Robust standard errors were calculated by clustering on these 88 groups.

Extended Data Table 5 | Experiment 2, the effects of the treatments on the movie-goers for subjects with a baseline implicit association score at or below the median. The response variable consists of implicit association scores normalised to have a mean of zero and a standard deviation of one. Estimates show the difference-in-difference effects by treatment relative to the control (Supplementary Information, § 4). To check for robustness, we analysed the unweighted complete cases, the inverse probability weighting of complete cases, and multiple imputations. Data include 6659 observations from 3541 participants. Under multiple imputation, we imputed 423 observations for each of 200 imputations. For treatment assignment, 122 communities were combined into 88 groups according to geographic proximity. Robust standard errors were calculated by clustering on these 88 groups.

Extended Data Table 6 | Experiment 2, intention-to-treat effects for subjects with a baseline implicit association score above the median. The response variable consists of implicit association scores normalised to have a mean of zero and a standard deviation of one. Interactions between the follow-up data collection and the treatment dummies (e.g. Follow-up \times combined; Supplementary Information, § 4) estimate the difference-in-difference effects relative to the control. These effects are highlighted in blue. To ensure results robust to missing data, we analysed the unweighted complete cases, the inverse probability weighting of complete cases, and multiple imputations. Data include 6647 observations from 3541 participants. Under multiple imputation, we imputed 435 observations for each of 200 imputations. For treatment assignment, 122 communities were combined into 88 groups according to geographic proximity. Robust standard errors were calculated by clustering on these 88 groups.

Extended Data Table 7 | Experiment 2, the effects of the treatments on the movie-goers for subjects with a baseline implicit association score above the median. The response variable consists of implicit association scores normalised to have a mean of zero and a standard deviation of one. Estimates show the difference-in-difference effects by treatment relative to the control (Supplementary Information, § 4). To check for robustness, we analysed the unweighted complete cases, the inverse probability weighting of complete cases, and multiple imputations. Data include 6647 observations from 3541 participants. Under multiple imputation, we imputed 435 observations for each of 200 imputations. For treatment assignment, 122 communities were combined into 88 groups according to geographic proximity. Robust standard errors were calculated by clustering on these 88 groups.

Methods

In the **values** movie (Supplementary Information, § 2), characters present arguments based on their individual evaluations of cutting. For some, health is paramount. They want to avoid a decision that could lead to health problems for their daughters. Others worry that Islam requires cutting. Still others fear that, if they do not cut, their daughters will not grow up to be feminine, morally upright women. These different views share the feature that they are all based on individual values that do not explicitly depend on whether other families cut¹⁵.

Altogether, the values movie begins by depicting a set of competing fears and concerns. The characters then examine and eventually alleviate the fears and concerns that favour cutting. For example, one day the mothers discuss a young woman, Samia, who is widely regarded for her good behaviour. The discussion revolves around the question, is Samia well behaved because she is cut or because of how her mother raised her? After considerable debate, the women conclude that Samia is well-behaved because of how her mother raised her. Additionally, both the mothers and fathers discuss whether Islam requires cutting. Again after considerable debate, two key points lead them to conclude that it does not. First, the oldest brother in the family lives in Saudi Arabia, and he has revealed that in Saudi Arabia, the birthplace of Islam, families do not cut. Moreover, in a crucial scene the mothers watch a Sheikh on television as he explains that Islam does not require cutting. To record this scene, we recruited a well-known Sheikh in Sudan who is an advocate for abandoning female genital cutting, and he played himself in the movie.

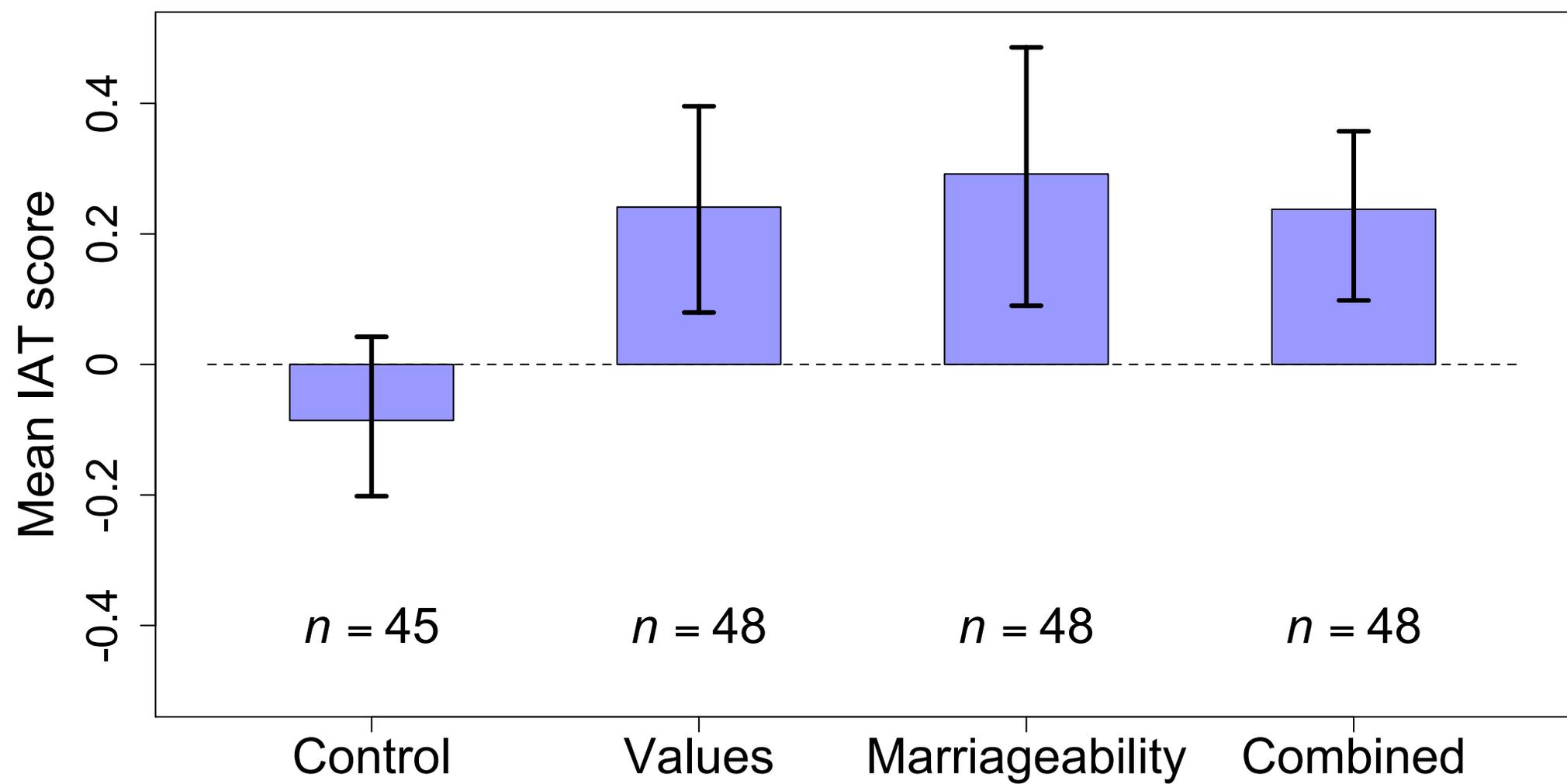
The **marriageability** movie (Supplementary Information, § 2) focuses on the question of

how cutting will affect the future marriage prospects of the daughters in the family. Specifically, everyone in the family agrees they must consider how cutting will influence whether their daughters grow up to find good husbands. Family members disagree, however, about the details. Some argue that times are changing, and cutting will soon disappear. If the family waits to abandon cutting, they will be viewed as backward, and their daughters will pay the cost. Others concede that some families may have stopped cutting, but they argue that change, if it comes at all, will take generations. Thus, the best way to guarantee a good future for their daughters right now is to cut them.

The dispute is resolved as family members gather more information. They learn that other families in the area have publicly abandoned cutting, and a record of this is available in the local community office. The family members also discuss the possibility that, if they abandon cutting, the decision might hasten the decision of other families to stop because these other families are also trying to anticipate the trend. In addition, the family considers that, even if they abandon cutting before most other families, their daughters can still marry within the extended family. Extended families are a common source of marriage partners in Sudan, and distant relatives within extended families often marry. Accordingly, the family recognises that, if they stop cutting together, the extended family itself will form an interim marriage pool for their daughters until cutting is abandoned more widely. Finally, in a crucial scene the mothers watch a public declaration of abandonment on television. In this scene, several families on television publicly declare that they have abandoned cutting, and by doing so all at once they collectively ensure good marriage prospects for their uncut daughters in the future.

The **combined** movie (Supplementary Information, § 2) integrates scenes and dialogues from both the values movie and the marriageability movie. In this sense it combines both heterogeneous values related to cutting and heterogeneous beliefs about how cutting will affect future marriage prospects. The **control** movie consists only of the main plot (Supplementary Information, § 2). All movies are approximately 90 minutes long. Subplots about cutting constitute approximately 30% of this (c. 27 minutes). We extended several scenes in the control movie to make the total length 90 minutes. To do so, we did not change the storyline or narrative structure of the main plot in any way.

We conducted both experiments in the state of Gezira, Sudan (Supplementary Information, § 3-4). We used different communities for the two studies. Adult participants came from randomly selected households with a roughly equal mix of men and women in each community. For the second experiment, we collected data with recruited participants, but each movie screening was open to the entire community. The studies were approved by the National Council for Child Welfare in Khartoum, the Gezira State Council for Child Welfare in Wad Madani, the Gezira State Ministries of Health and Education, all relevant authorities in all 127 communities, and the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology at the University of Zurich. Participants provided verbal consent after being briefed about the study.

ω  σ 