Article

The risk of female genital cutting in Europe: Comparing immigrant attitudes toward uncut girls with attitudes in a practicing country

Sonja Vogt*, Charles Efferson, Ernst Fehr

Department of Economics, University of Zurich, Blumlisalpstrasse 10, 8006 Zurich Switzerland

ARTICLE INFO

ABSTRACT

Worldwide, an estimated 200 million girls and women have been subjected to female genital cutting. Female genital cutting is defined as an intentional injury to the female genitalia without medical justification. The practice occurs in at least 29 countries in Africa, the Middle East, and Asia. In addition, globalization and migration have brought immigrants from countries where cutting is commonly practiced to countries where cutting is not traditionally practiced and may even be illegal. In countries receiving immigrants, governments and development agencies would like to know if girls with parents who immigrated from practicing countries are at risk of being cut. Risk assessments, for example, could help governments identify the need for programs promoting the abandonment of cutting among immigrants. Extrapolating from the prevalence and incidence rates in practicing countries, however, is generally not sufficient to guarantee a valid estimate of risk in immigrant populations. In particular, immigrants might differ from their counterparts in the country of origin in terms of attitudes toward female genital cutting. Attitudes can differ because migrants represent a special sample of people from the country of origin or because immigrants acculturate after arriving in a new country. To examine these possibilities, we used a fully anonymous, computerized task to elicit implicit attitudes toward female genital cutting among Sudanese immigrants living in Switzerland and Sudanese people in Sudan. Results show that Sudanese immigrants in Switzerland were significantly more positive about uncut girls than Sudanese in Sudan, and that selective migration out of Sudan likely contributed substantially to this difference. We conclude by suggesting how our method could potentially be coupled with recent efforts to refine extrapolation methods for estimating cutting risk among immigrant populations. More broadly, our results highlight the need to better understand how heterogeneous attitudes can affect the risk of cutting among immigrant communities and in countries of origin.

1. Introduction

Apart from the estimated 200 million girls and women living with the consequences of female genital cutting, an additional three million girls are at risk of being cut every year (World Health Organization, 2014). Cutting is historically practiced in several countries in Africa, Asia, and the Middle East. As globalization proceeds apace, however, governments and international organizations are increasingly concerned about the prevalence and risk of cutting among immigrant populations in Europe and North America. Cutting among immigrants is important for a number of reasons. Immigrants who have been cut, for example, may have special needs in terms of health care. They may find it difficult or impossible to get the care they need if they live where doctors have little or no experience with patients who have been cut and possibly infibulated. Moreover, policy makers also need to understand the risk of being cut for girls with parents who have moved from countries with a history of cutting to other countries where cutting is

uncommon and even illegal. This risk will likely determine the extent to which governments invest in efforts to promote the abandonment of cutting among immigrants. For these reasons, the European parliament has called for better data and better methods to estimate the number of women and girls cut or at risk of being cut in Europe (European Institute for Gender Equality, 2013, 2016).

Prevalence in an immigrant population refers to the proportion of girls and women who migrated from the associated practicing country and have already been cut (European Institute for Gender Equality, 2013). Girls at risk of cutting refers to uncut girls aged 18 or younger who migrated from a practicing country or have at least one parent who migrated from a practicing country (European Institute for Gender Equality, 2016). Most estimates of prevalence and risk in immigrant populations are derived by extrapolating from prevalence data in countries of origin (European Institute for Gender Equality, 2013). These data typically come from representative surveys, including the Demographic Health Survey and the Multiple Indicator Cluster Survey,

---

* Corresponding author.

periodically implemented in the 29 or more countries where female genital cutting is traditionally practiced (Yoder, Wang, & Johansen, 2013). The simplest approach to extrapolation is to apply an estimate of prevalence from the country of origin to an immigrant population. For example, if an estimated 80% of Sudanese women in Sudan are cut, then an estimated 80% of female Sudanese immigrants in Switzerland are cut. Recently, however, a number of researchers have developed refinements that extrapolate while accounting for differences between an immigrant population and people in the associated country of origin (Exterkate, 2013; Leye, Mergaert, Arnaut, & Green, 2014). These refined methods consider, for instance, age-specific cutting rates, regional differences in cutting within a country of origin, or differences between first-generation and second-generation immigrants (Dubourg et al., 2011; Exterkate, 2013; Leye et al., 2014; Ziyada, Norberg-Schulz, & Johansen, 2016).

These refinements have the potential to greatly improve estimates of risk and prevalence because they allow for the possibility that immigrants are not necessarily like the people who remained in the country of origin. For example, one of the challenges when estimating risk among immigrants is that the age of cutting varies greatly within and between practicing countries. Ignoring this fact can distort estimated risk considerably. To illustrate, assume that every girl in a country of origin is cut the day she turns five. For every female aged five or older, prevalence is thus 100%. Further assume that in Europe all female immigrants from this country immigrated before they turned five, and they are all now above five. If we ignore these subtleties, we might conclude that the prevalence among immigrants in Europe is 100%, when in reality it could be 0% because everyone arrived in Europe before the cutting age. Dubourg et al. (2011), Exterkate (2013), and Ziyada et al. (2016) accounted for subtleties of this sort with extrapolation methods that condition on age-specific prevalence in countries of origin.

Nonetheless, even if we disaggregate prevalence data from countries of origin based on observables like age and region, extrapolation may still not fully account for selective emigration from a cutting country, and it cannot account for acculturation after immigration into a non-cutting country (Leye et al., 2014). For this reason, researchers have developed the "migration and acculturation impact factor" (Exterkate, 2013; European Institute for Gender Equality, 2016; Ortensi et al., 2015). The general idea behind this factor is to refine estimates of risk for immigrants by jointly accounting for selective migration out of the country of origin and for cultural and behavioral changes after immigrating to a new country (European Institute for Gender Equality, 2016). For example, if everyone in a particular immigrant group comes from a non-cutting area in the country of origin, the risk is zero, regardless of aggregate risk and prevalence in the country of origin. Analogously, if everyone in a particular immigrant population completely and immediately assimilates to the norms of their new home, the risk that any uncut immigrant will be cut is zero, regardless of aggregate risk and prevalence in the country of origin. Selective migration and acculturation are two very different but potentially important channels that can lead the risk of cutting among immigrants to diverge from the country of origin.

The present study directly compares attitudes about female genital cutting among Sudanese living in Switzerland to attitudes among Sudanese living in Sudan. Whether or not Sudanese immigrants are a special subset of the Sudanese population, and regardless of whether or not they move toward the values of Switzerland after immigrating, we focus directly on identifying any attitudinal differences by using the same fully anonymous method to measure implicit attitudes in both countries. Importantly, we previously validated our implicit attitudinal measure by showing that measured attitudes in Sudan were highly correlated with incidence at the community level (Vogt, Zaid, Ahmed, Fehr, & Efferson, 2016). This correlation provides us with a relationship between attitudes and a girl's risk of being cut. We use this relationship in conjunction with our attitudinal data from Switzerland

to discuss the magnitude of the migration and acculturation impact factor on the risk of being cut for girls in the Sudanese immigrant population. Moreover, using control data collected in Switzerland, we also provide results relevant to the question of whether attitudinal differences are due to selective migration or acculturation.

Finally, the methods presented here also suggest how researchers might develop robust attitudinal measures for evaluating programs that promote the abandonment of cutting among immigrants. Because all extrapolation methods begin with data from a country of origin, extrapolation can never provide a basis for evaluating programs targeted specifically at immigrants. Evaluating a program that promotes abandonment in Europe, for example, requires validated methods that can be used among immigrants themselves.

2. Effects of selection and acculturation

The extrapolation method is a feasible and practical method for the difficult problem of estimating the risk of cutting among immigrant girls (European Institute for Gender Equality, 2013, 2016). It has the particular advantage that it relies on representative surveys in countries of origin. Representative samples among immigrant populations are extremely challenging. In general, we do not have country-wide sampling frames that include first-generation and second-generation immigrants, families with only one parent from a country where cutting is practiced, and asylum seekers who are not yet registered (Ziyada et al., 2016). In addition to the fact that we often do not know how to delimit the immigrant population of interest, we also have little understanding of how selective migration and acculturation affect attitudes toward female genital cutting (European Institute for Gender Equality, 2016).

First, migrants may constitute a special subset of the population of origin (Leye et al., 2014; Ortensi et al., 2015). Research has demonstrated that attitudes about cutting in many African countries vary tremendously among and even within households in a local area (Bellemare, Novak, & Steinmetz, 2015; Efferson, Vogt, Elhadi, Ahmed, & Fehr, 2015; Hernlund & Shell-Duncan, 2007). This kind of variation implies ample scope for emigration that is somehow conditional on attitudes about cutting (Ortensi, Farina, & Menonna, 2015). If this kind of selection occurs, those who emigrate will not have the same attitudes regarding cutting as those who do not. Farina and Ortensi (2014), for example, surveyed immigrants from practicing countries in Italy and concluded that ignoring selective migration can lead one to overestimate the prevalence and risk of cutting in immigrant populations. Second, in addition to selection, immigrants in Europe and North America might have attitudes about cutting that differ from countries of origin because immigrants change their attitudes after arriving in their new country (Farina & Ortensi, 2014; Johnsdotter et al., 2009). Some might become more negative about cutting as they integrate in a non-cutting society, while others might become more positive about cutting as a way to maintain and assert cultural ties to their native countries. All in all we know little about the net effect of attitudinal changes after migrating (Leye et al., 2014). Generically, however, we expect both selective migration and attitudinal changes to considerably complicate the task of estimating risk among immigrants.

Recent extrapolation studies have confronted this problem by developing a number of techniques for adjusting risk based on the joint effect of selection and acculturation. Exterkate (2013) used insights from focus groups to specify high-risk, medium-risk, and low-risk scenarios for immigrant girls in the Netherlands. Ortensi et al. (2015) addressed selection by considering variation in prevalence among regions within practicing countries. They specifically distinguished between countries where cutting is widespread and countries where the practice is regionally clustered, and they used this information to adjust risk estimates for immigrant populations. Ziyada et al. (2016) carefully distinguished between the cutting risk among first-

generation and second-generation immigrants, and they additionally accounted for the age of immigration. More broadly, the European Institute for Gender Equality (2016) has produced a detailed and thorough review of how to address the potential effects of selective migration and acculturation.

In our case, we test if attitudes toward cut and uncut girls differ between Sudanese people living in Switzerland versus Sudan. This is a test of the joint effect of selection and acculturation. We additionally analyze variation in attitudes within Switzerland, including an analysis of whether attitudes among immigrants depend on how long an immigrant has been in Switzerland and other measures of integration. This analysis specifically addresses the effects of acculturation (Exterkate, 2013; Ortensi et al., 2015).

3. Anonymous implicit measures

Examining attitudes about female genital cutting requires us to measure attitudes in some way. Studies in practicing countries have been largely based on measuring attitudes and cutting behavior via face-to-face interviews. As one example, interviewers for the Demographic Health Survey ask women directly whether they are cut and whether they think cutting should continue (Yoder & Khan, 2008; Yoder et al., 2013). Studies surveying immigrant attitudes are also largely based on face-to-face interviews or focus groups, and these studies have provided critical comparative data and important insights about female genital cutting among immigrants (Gele, Johansen, & Sundby, 2012, Gele, Sagbakken, & Kumar, 2015; Isman, Ekeus, & Bergen, 2013; Johnsdotter, Moussa, Carlbom, Aregai, & Essen, 2009; Morison, Dirir, Elmi, Warsame, & Dirir, 2014; Reig Alcaraz, Siles González, & Solano Ruiz, 2014).[1] Morison et al. (2014) showed, for instance, that immigrating at a younger age reduced support for female genital cutting, and Gele et al. (2012) found that migrants supported cutting less the longer they had lived in a non-cutting country.

Face-to-face surveys typically use explicit measures of attitudes. With potentially sensitive topics like female genital cutting, an explicit approach can lead to biased answers (Gruenbaum, 2005; Mackie, 2000). Respondents may under-report their support for cutting (Blattman, Jamison, Koroknay-Palicz, Rodrigues, & Sheridan, 2015; Glynn et al., 2011; Schroder, Carey, & Vanable, 2003) due to social desirability biases, especially if respondents have been recently exposed to an intervention promoting the abandonment of cutting (de Cao & Lutz, 2015). On the other hand, respondents in some places might over-report their support for cutting as a way to assert their cultural autonomy in the face of unwelcome intrusions from governmental and non-governmental workers known to favor abandonment. This scenario would be consistent with the backlashes that have sometimes occurred against programs promoting abandonment (Boddy, 2007; Camilotti, 2015; Gruenbaum, 1996, 2001; Shell-Duncan & Hernlund, 2000; Shell-Duncan, Wander, Hernlund, & Moreau, 2013; Thomas, 2000).

The net effect and geographic variability of under-reporting and over-reporting is completely unknown, and by extension extrapolation could be unreliable because it depends on data collected with explicit methods that are quite malleable and thus susceptible to bias. In particular, to the extent that explicit measures allow respondents to strategically manipulate their responses easily, a comparison of data between migrants and non-migrants could be especially fraught with problems. If cutting is illegal in the immigrant country but not in the country of origin, or if there are differences in the enforcement of the law, immigrants may systematically under-report their support for

cutting, while residents in the country of origin may feel no need to do so.

To limit the risk of such problems, the current study compares implicit attitudes on cut versus uncut girls using data collected under fully anonymous conditions. Anonymity can dramatically reduce biases when collecting data on sensitive issues (Chauchard, 2013), and implicit measures can further reduce biases (de Cao & Lutz, 2015; Greenwald et al., 2003; Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Nosek, Greenwald, & Banaji, 2007). For these reasons, we use a validated measure of implicit attitudes to collect data under anonymous conditions and compare Sudanese immigrants in Switzerland with non-migrants in Sudan.

4. Material and methods

4.1. Implicit attitudes on female genital cutting

We developed, validated, and implemented an implicit association test to measure attitudes about female genital cutting among people of Sudanese origin in both Switzerland and Sudan. The test was fully computerized. It did not require participants to read and write, and it did not require participants to have any prior experience with computers. Participants simply listened to audio instructions over headphones, and during the test they used only three keys, a key with a pink sticker on the left side of the keyboard, a key with a blue sticker on the right side, and the space bar highlighted with a white sticker.

Implicit association tests have been widely used in social psychology, neuroscience, and development economics (Beaman, Chattopadhyay, Duflo, Pande, & Topalova, 2009; Nosek et al., 2007; Rooth, 2007). Such tests are designed specifically to minimize the potential for respondents to manipulate their answers. They require a respondent to categorize target stimuli and valued stimuli as quickly as possible. In our case, the target stimuli were various drawings of two girls who were introduced to the respondent as cut or not cut (Fig. 1). We used two different dresses as mnemonic devices to signal which girl was cut and which was not. Specifically, the two girls themselves were distinguished by the fact that one had pigtails, while the other did not. Aside from this, one girl wore a "firka" dress, and the other girl wore a "saleema" dress. Firka is a fabric with a distinctive and instantly recognizable pattern. The fabric plays a prominent role in the cutting ceremony in Sudan, and it has a strong association with cutting. For this reason, the cut girl wore a dress made from firka (Fig. 1). The uncut girl, in contrast, wore a dress made from the saleema fabric. This fabric is associated with the saleema campaign in Sudan, which is largely a media-based campaign promoting the abandonment of cutting. However, the saleema campaign had not yet run at our study site in Sudan at the time of our study.

The valued stimuli were audio recordings of positive words (great, good, amazing, all right, beautiful) and negative words (worthless, ugly, unimportant, bad, disgusting). Thus, altogether we had four stimulus categories. A given stimulus was either a drawing of the cut girl in the firka dress, a drawing of the uncut girl in the saleema dress, an audio recording of a positive word, or an audio recording of a negative word. For the test itself, stimuli appeared rapidly on the computer screen one at a time. The respondent's task was to categorize each stimulus by pressing one of two keys on the computer. This is the central feature to the implicit association test. Each respondent was presented with stimuli in four different categories, and the respondent had to reduce this to two categories. The speed with which the respondent did so was the basic response variable. Each participant had to sort stimuli under two different categorizations schemes, and her relative speed under these two different schemes provides a measure of the participant's implicit attitudes.

One categorization scheme paired drawings of the cut girl with negative words and drawings of the uncut girl with positive words (Fig. 2A). This means, if the respondent saw either a drawing of the cut

---

[1] Other studies have focused on interviewing health-care professionals (Hänselmann, Börsch, Ikenberg, Strehlau, & Klug, 2011; Jager, Schulze, & Hohlfeld, 2002; Korfker et al., 2012). These studies have highlighted the need for special forms of health care among immigrants, and they have raised important questions about whether health care workers are trained to handle cases related to cutting.

Fig. 1. On the left side is the picture of the girl without pigtails in a dress made from the firka cloth. The girl is introduced as recently cut with parents who cut all their daughters. On the right side is the picture of the girl with pigtails in a dress made from the saleema cloth. The girl is introduced as uncut, and she will not be cut as her parents do not cut any of their daughters. Reprinted with permission from UNICEF, Sudan.
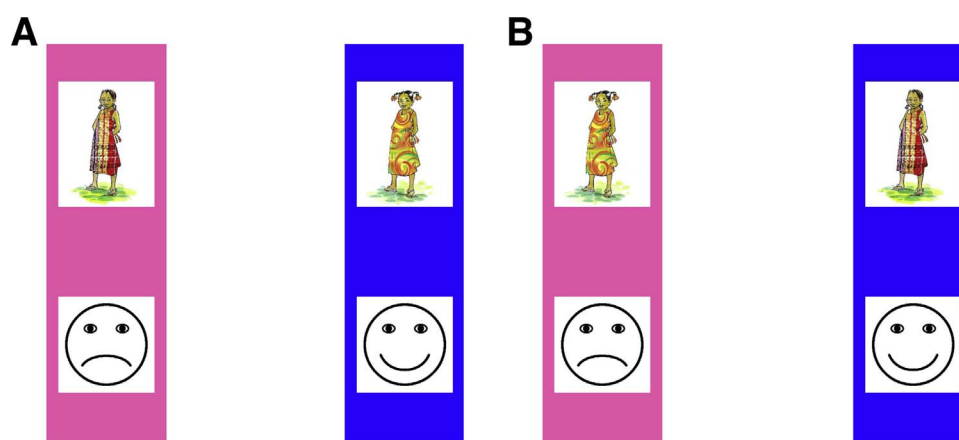

Fig. 2. The two different categorization schemes for the implicit association test. The panels shown were displayed on the computer screen to remind respondents of the current scheme. Each stimulus was either a drawing of a girl that appeared suddenly in the middle of the screen or an audio recording of a word. Under the categorization scheme in panel **A**, if a respondent saw a drawing of the cut girl or heard a negative word, she had to press the pink key on the left side of the keyboard as quickly as possible. If the respondent saw a drawing of the uncut girl or heard a positive word, she had to press the blue key on the right side of the keyboard as quickly as possible. The frowning face and the smiling face were used to represent negative and positive words respectively. Under the categorization scheme in panel **B**, the associations were exactly the opposite, with uncut paired with negative, and cut paired with positive. See Fig. 3 for example screens with drawings of girls. Reprinted with permission from UNICEF, Sudan. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

girl or heard a negative word, she had to press a pink key on the left side of the keyboard as quickly as possible. Alternatively, if a respondent saw either a drawing of the uncut girl or heard a positive word, she had to press a blue key on the right side of the keyboard as quickly as possible. The other categorization scheme simply reversed the pairing (Fig. 2B). If the respondent saw either a drawing of the uncut girl or heard a negative word, she had to press the pink key on the left. If the respondent saw either a drawing of the cut girl or heard a positive word, she had to press the blue key on the right. Fig. 3 shows example screens with drawings of the two girls presented as stimuli.

The logic behind the test is that response times are quicker on average for the categorization scheme consistent with the respondent's implicit attitudes. If a given respondent is pro-cutting, for example, she will tend to be relatively fast when cut is paired with positive and uncut with negative, and she will tend to be relatively slow when cut is paired with negative and uncut with positive. Across participants, we used an extensive counterbalancing scheme to eliminate the possibility of any biases or artifacts unrelated to actual implicit attitudes about cutting. In particular, we used counterbalancing to ensure that approximately half of our participants took a version of the implicit association test in
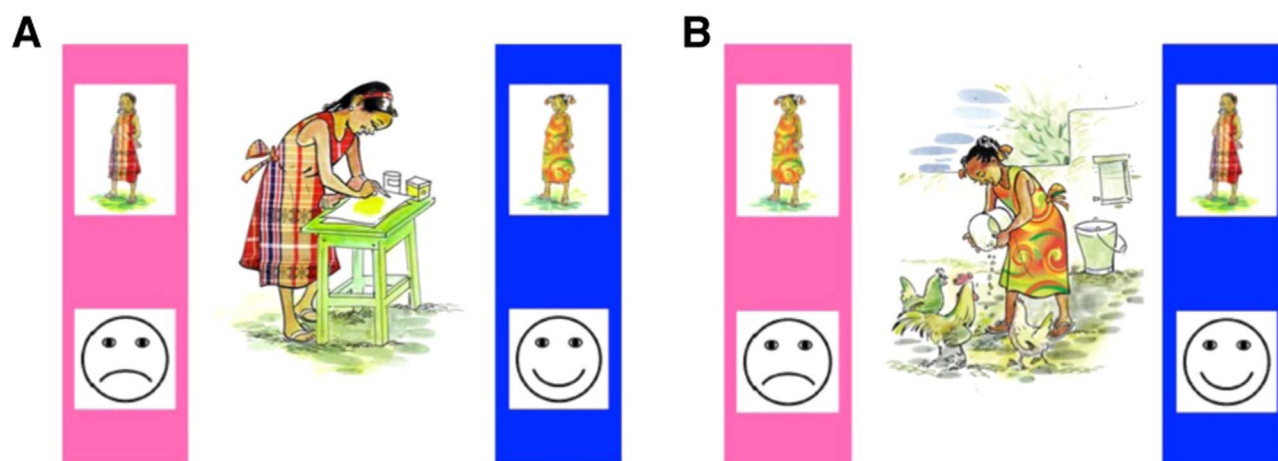
Fig. 3. Panel **A** shows an example in which a drawing of the cut girl appears in the middle of the screen, and the correct response is to press the pink key on the left as quickly as possible. Panel **B** shows an example in which a drawing of the uncut girl appears in the middle of the screen. Because the categorization scheme in **B** is the opposite of that in **A**, the correct response is again to press the pink key on the left as quickly as possible. Reprinted with permission from UNICEF, Sudan. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

which the girl without the pigtails was cut, while the other half took a version in which the girl with the pigtails was cut. We also counter-balanced the order of the categorization rules across participants. Thus, approximately half of our participants started with the categorization scheme that paired the uncut girl with negative words, while the remaining participants started with the categorization scheme that paired the cut girl with negative words. Altogether, we thus had four different versions of our implicit association test, and our counter-balancing strategy ensured that they were randomly and uniformly distributed among our participants. More details are available in Vogt et al. (2016), (see Supplementary material).

Relative response times under the two different categorization schemes allow us to calculate a summary measure of implicit associa-tions. This measure is the D statistic for each participant, and it is essentially a normalized measure of relative response times under the two categorization schemes (Nosek et al., 2007). D is normalized such that $D \in [-2, 2]$. $D \in [-2, 0)$ indicates relatively negative implicit associations with uncut girls, while $D \in (0, 2]$ indicates relatively positive implicit associations with uncut girls. $D = 0$ indicates the participant had no implicit associations. The normalization used to calculate D for each participant accounts for heterogeneity in average response times across participants and the variance in response times for a given participant. Thus, D is a valid metric even if participants differ in terms of how long they take and in terms of how variable each participant's response times are across trials.

In a previous study (Efferson et al., 2015), we implemented our implicit association test with 2260 randomly selected adults in 45 communities in Sudan. In addition, we also estimated the incidence of cutting among girls entering primary school in the same 45 commu-nities using a new method that did not require us to ask parents directly about their cutting practices. This method was based on the fact that young girls in our study site get henna on their feet on the day they are cut. Henna is almost always applied to a young girl's feet when she is cut, and this is the only time when henna is applied to a young girl's feet. Residual henna remains on the toenails for several weeks. As girls are usually cut in the summer vacation shortly before beginning primary school or earlier, we implemented the following procedure with all girls entering primary school. Over the course of the first few days of the school year, nine teams of photographers and medical doctors visited the schools in all 45 communities in the study. The photographers photographed the feet of all girls entering school, and the doctors administered a basic medical exam. As part of the medical exam, doctors asked the girls if they had been "purified," which is a translation of the local term used for cutting. We constructed a measure of the incidence of cutting in each community by combining data from the photographs and from the responses to questions asked during the medical exam (Efferson et al., 2015, see Supplementary material). To preserve the complete anonymity of participants, both the adults who took the implicit association test and the girls entering primary school, our data from this study included no identifying information of any kind at the individual level. By combining the two types of data, we found a highly significant correlation between average implicit association scores and incidence values by community (Fig. 4).

4.2. Study approval

The Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology at the University of Zurich approved the studies in Switzerland and Sudan. In addition, the Sudanese National Council for Child Welfare, the Gezira State Council for Child Welfare, the Gezira Ministries of Health and Education, and all relevant community authorities in all commu-nities approved the study in Sudan. Sudanese assistants handled recruitment and data collection in both Sudan and Switzerland. Participation was strictly voluntary and based on informed verbal
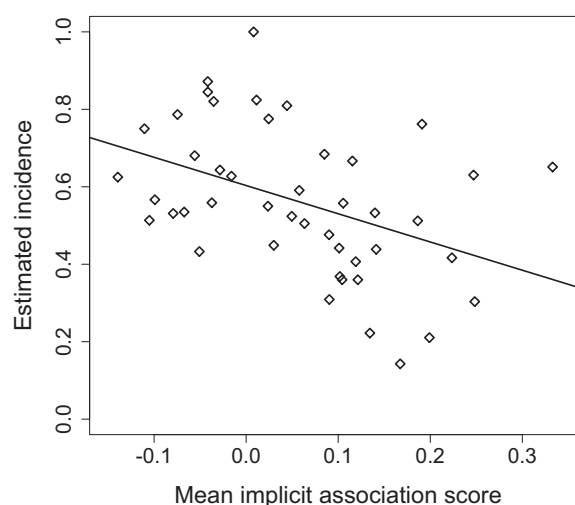

Fig. 4. The relationship between average implicit association scores by community and estimated incidence across 45 communities in Gezira, Sudan. The line shows the unweighted least squares line. The correlation is highly significant (Pearson's correlation, $\rho = 0.423$, one-sided $p=0.008$ based on two-dimensional weighted bootstrap explained in the Supplementary material for Vogt et al. (2016)).

consent. Data collection was anonymous, and participants were rewarded for their participation. Below we explain the recruitment of participants and data collection for Sudan and Switzerland.

4.3. Recruitment and procedures for the implicit association test in Sudan

In a previous study (Efferson et al., 2015), we elicited implicit attitudes about cutting among 2260 randomly selected adults in 45 communities in the state of Gezira, Sudan. In particular, community leaders maintain lists of households and household members in their communities. After reviewing and updating these lists with community leaders, we randomly sampled households in each community. After sampling households, we contacted each household individually and recruited a single adult participant. In half of the sampled households in a community, we recruited an adult female. In the other half, we recruited an adult male. Random sampling led to a wide variety of participants from different tribes. Altogether, based on their own self-reported affiliations, the 2260 adults from the 45 communities represent 60 different sub-tribes, which can be grouped into 31 main tribes.[2]

We focused on a single state in Sudan to facilitate data collection given several logistic and bureaucratic challenges. In particular, we had four teams of data collectors, each consisting of approximately 10–15 facilitators. These teams traveled around Gezira collecting data on a daily basis with all the materials necessary to set up and use a mobile computer lab with 10–12 computers in rural areas. In addition, because our participants were both men and women, we needed facilitators who were both men and women. To do this in Sudan, a country with cultural restrictions on the movement of women, we needed to ensure that our female facilitators would be home at the end of each day. This meant we had to restrict one-way travel time from Wad Madani, the capital of Gezira, to two to three hours. Gezira has relatively good roads, and for this reason our facilitators could work in many communities but still get home at the end of the day. In addition, we required the approval of local government and community leaders in all communities. This was a challenge to organize and would have been virtually impossible to do so in a consistent fashion across Sudan. Given these constraints, we chose Gezira for at least three important reasons. First, rural communities in Gezira have *not* been extensively exposed to campaigns promoting the abandonment of female genital cutting. In addition to our use of implicit attitudinal measures under anonymous conditions, this should further reduce the scope for biased responses (de Cao & Lutz, 2015). Second, as mentioned, the roads in Gezira allowed our facilitators to work their way through a relatively large geographic area and still get home at the end of the day. Finally, representative survey data show that Gezira is nearly identical to Sudan as a whole in terms of numerous indicators related to health, education, hygiene, and the prevalence of cutting (Central Bureau of Statistics (CBS), 2016). Therefore, although we cannot categorically guarantee that our data from Gezira are representative of Sudan more broadly, existing data suggest that Gezira is not an unusual state and should thus provide a good benchmark in terms of attitudes and practices related to cutting.

Data collection took place primarily in community school buildings. For a given session, facilitators built a mobile computer lab in the school building, and each computer station was surrounded by a wooden partition to protect the participant's privacy during the implicit association test. Each participant began by working with a facilitator. The facilitator first taught the participant how to use the pink and blue keys with an abbreviated implicit association test unrelated to female

genital cutting. This abbreviated test used the bugs vs. flowers stimuli from Baron & Banaji (2006).[3] They repeated this exercise if necessary. After this introduction, the facilitator would start the implicit association test on female genital cutting. The facilitator would help the participant to put on headphones and then start the actual test. The facilitator would do a quick audio check with the headphones and then immediately leave the participant to complete the test in privacy. However, the facilitator would remain within view of the participant in case problems arose. Participation was reimbursed with prayer rugs, perfume, or other household items of similar value. See the electronic supplements for Efferson et al. (2015) and Vogt et al. (2016) to find more details.

4.4. Recruitment and procedures for the implicit association test in Switzerland

We worked with Sudanese adults currently living in Switzerland. According to the Swiss Federal Statistical Office, far fewer than 1000 adults of Sudanese origin live in Switzerland (Swiss Federal Statistical Office, 2016). We did not constrain participation in any way based on residency status, how long a person had been living in Switzerland, or how long the person lived in Sudan. However, participants had to be fluent in colloquial Sudanese Arabic. Altogether 84 participants participated in Switzerland. About 30% were female, 88% were born in Sudan, and the average age was 35.48 years (18–82, $\sigma = 12.18$). The average time in Switzerland was 11.1 years (2 months to 55 years, $\sigma = 9.99$ years).

Existing studies on female genital cutting in immigrant populations often rely on respondent-driven sampling techniques (Gele et al., 2012; Isman et al., 2013; Johnsdotter et al., 2009; Morison et al., 2014; Reig Alcaraz et al., 2014). Such techniques are often convenient because a well-defined sampling frame does not exist for many immigrant communities. In addition, gaining access to immigrant communities with the intention of studying a highly sensitive topic like female genital cutting can be quite challenging, especially if cutting is criminalized as it is in Switzerland. Respondent-driven sampling methods use network ties within the population of interest as a way around these problems. The downside, however, is that such methods can exaggerate selection bias. If a recruited participant helps recruit her friends for participation, the researcher can end up with a relatively homogeneous and unrepresentative sample of like-minded people. In our case, our primary objective was to reduce these problems as much as possible given that we did not have, as is often the case, a sampling frame. Accordingly, we used the following approach.

We conducted an initial study that was highly attractive from the perspective of participants, and it had nothing whatsoever to do with female genital cutting. The Swiss Sudanese Association in Zurich invited us to run this initial study during their Eid festival. Although this festival takes place in Zurich, members are widely dispersed. Indeed, the festival is one of very few large events in which Sudanese families distributed around Switzerland come together each year. The study involved an incentivized behavioral experiment on altruism. Specifically, in cooperation with the Sudanese association, we set up a mobile computer laboratory consisting of over 30 computers at the festival. People at the festival had the opportunity to join our study spontaneously. Although the study only required a few minutes to complete, it offered a sizable maximum reward of 50 Swiss Francs. The amount of money a person actually earned depended on the decision she made during the experiment. For this initial study, we had no information of any kind on our participants. We did not know their names, and we had no contact information.

Nonetheless, after the study any adult at the festival could sign up for participation in future studies by giving us only a name and phone

---

[2] We would like to thank Dr. Mohamed Abdalla Mohamed, a local physician intimately familiar with all the communities in our study area, for re-grouping the sub-tribes into main tribes.

[3] We would like to thank Mahzarin Banaji for providing us with these materials.
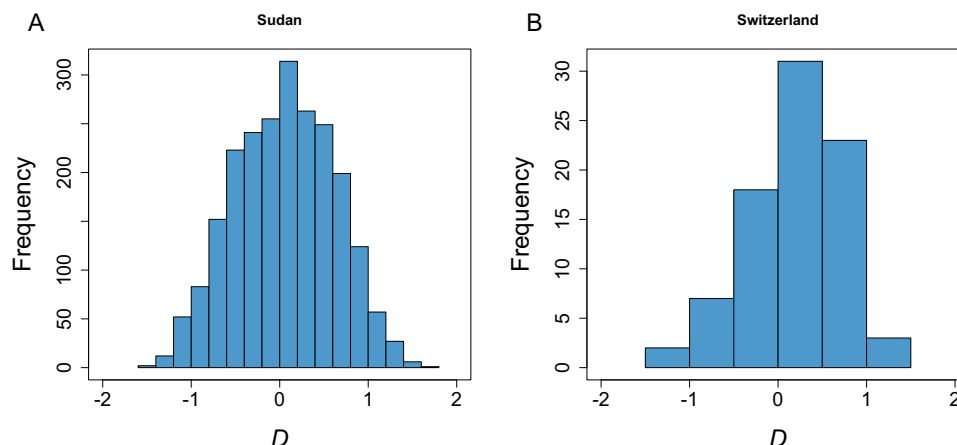
Fig. 5. The graph shows the distribution of implicit association scores (D) over all 2260 participants in Sudan (**A**, $\overline{D} = 0.058$) and the distribution of scores over all 84 participants in Switzerland (**B**, $\overline{D} = 0.201$). The two scores are significantly different (Wilcoxon rank sum test, two-sided, $p=0.0105$).

number. In the final analysis, given that this initial experiment only required a few minutes to complete and offered a considerable reward, the vast majority of the adults at the festival participated and provided us with contact details for participation in future experiments. In addition, we later received a list of members from the Swiss Sudanese Association in Zurich with allowance to contact them for future studies. Together these two lists of potential participants provided a basis for recruiting participants for the present study on female genital cutting.

For the present study, we contacted potential participants by phone. We told them at that time that the study would concern cutting. We also told them that the data from the study would be completely anonymous and that participation was completely voluntary. Finally, we told potential participants that each person who chose to participate would receive 50 Swiss Francs as compensation for their time.

We hired two assistants, one Swiss-Sudanese and the other a recent immigrant from Sudan, to recruit subjects, travel to their homes, and actually implement the implicit association tests. Although the facilitators visited the homes of participants for the study, the participants actually completed the implicit association test alone. The facilitators followed exactly the same procedure as described for our data collection in Sudan. Thus, participants first completed the same abbreviated implicit association test unrelated to female genital cutting with a facilitator present. After this, participants completed the implicit association test on female genital cutting in privacy. The data were fully anonymous. The facilitators had no way to analyze or interpret the data, and indeed response times were simply recorded by the computers as one among many variables stored in large electronic data files. When we received the data, we had only identification numbers for participants with no way to link a given number to a person.

Our recruitment procedure offers potential advantages when studying a culturally sensitive topic in a relatively inaccessible study population. In particular, because we developed our initial sampling frame with a behavioral experiment completely unrelated to female genital cutting, we were able to establish a favorable precedent with potential subjects. In particular, our initial experiment was fast, interesting, and it paid well. Most importantly, because this initial experiment did not address cutting in any way, we have no reason to believe that it could have generated a sampling frame of people who were, for example, systematically opposed to cutting. Moreover, our initial study highlighted to the community of Sudanese immigrants in Switzerland that we present them with tasks they complete in private, we secure the data in a way that further protects their privacy, and participation brings relatively large immediate rewards. Ultimately, although we did not manage to contact about 10% of the adults in our sampling frame, everyone we did contact agreed to participate.

The gender distribution in our current study is not representative for the Sudanese population in Switzerland. 45% of all Sudanese immigrants in Switzerland are female, but only 30% of participants in our study on implicit attitudes are female. To speculate about why we had more men in our sample than women, women might have been more likely to take care of children during the Eid festival, leaving more men free to participate in our initial study on altruism. However, to maximize the feeling of privacy at the Eid festival, we did not record any private data, including gender, during the behavioral experiment. Therefore, we cannot provide data to confirm our speculation that more men joined at the Eid festival than women. Although we cannot categorically exclude the possibility that the people who joined our initial study were somehow unusual in terms of attitudes about cutting, we can say that we made a considerable effort to limit selection biases of this sort.

5. Results

We begin by comparing implicit association scores among people of Sudanese origin living in Sudan to those living in Switzerland. We then examine systematic variation in implicit attitudes among Sudanese immigrants in Switzerland and discuss the results in terms of the relative importance of selective migration versus acculturation.

5.1. Comparing attitudes about cutting in Switzerland and Sudan

Recall that D is normalized such that $D \in [-2, 2]$. The average implicit association score in Sudan was 0.058, and the average score in Switzerland was 0.201. Sudanese immigrants in Switzerland have significantly more positive attitudes towards uncut girls than Sudanese in Sudan (Wilcoxon rank sum test, two-sided, $p=0.0105$). The variance in scores in Sudan was 0.316, and the variance in Switzerland was 0.245. These variances are significantly different (Brown-Forsythe Levene-type test, $p=0.048$). Fig. 5 shows the two distributions of implicit association scores for Sudan and Switzerland. The histograms show that the distribution from Switzerland is shifted to the right relative to the distribution from Sudan. Although D scores were significantly less variable in Switzerland than in Sudan, they nonetheless varied substantially in Switzerland, ranging all the way from strongly negative to strongly positive scores. We discuss this variation in attitudes at length in the discussion.

How can we interpret the magnitude of the difference in mean test scores? In Switzerland, the standard deviation in implicit association scores was $\sqrt{0.245} = 0.49$. By extension, the difference between Switzerland and Sudan, namely $0.201 - 0.058 = 0.143$, was approximately 29% of a standard deviation. This is a substantial difference. Moreover, approximately 59% of the scores in Sudan were below the

mean score in Switzerland, while only 41.7% of the scores in Switzerland were below the mean score in Switzerland.

Finally, we can interpret the difference in scores by evaluating what the difference could mean in a simple extrapolation exercise. Specifically, assume that the linear relationship between implicit attitudes and risk estimated in Sudan (Fig. 4) also holds in Switzerland. This relationship shows that the risk of cutting declines at a certain rate as implicit attitudes toward uncut girls become more favorable. Under this relationship, the implicit attitudes in Switzerland would translate into an estimated incidence of 0.457, as opposed to the incidence of 0.585 we found in Sudan. Importantly, we do not want to suggest that the risk of cutting in Switzerland is 0.457. The relationship between attitudes and incidence that we estimated in Sudan (Fig. 4) could be very different in Switzerland. In particular, given a specific implicit attitude, a Sudanese immigrant in Switzerland may be less likely to cut her daughter than an otherwise similar Sudanese parent in Sudan. With this important caveat, our task is simply to use the relationship estimated in Sudan as a way to gauge the difference in attitudes between Sudan and Switzerland. Holding the estimated relationship constant, the difference in attitudes would translate into a drop in incidence of almost 13% for Sudanese girls living in Switzerland.

5.2. Analyzing the variation in attitudes within Switzerland

In addition to implementing the implicit association test in Switzerland, we also collected basic demographic information from our participants. We specifically asked each participant's age, sex, if he or she had one or more daughters, and if he or she had one or more sons. We additionally collected data to measure each participant's level of integration. We asked participants how long they had been living in Switzerland, whether they were fluent in the local language, and whether they were employed. Time in Switzerland, fluency in the local language, and being employed should all increase exchange with families who do not practice cutting and exposure to Swiss norms and culture. These variables therefore serve as proxies for integration in Swiss society. Though of course not perfectly correlated with each other, the three variables are mutually consistent. Specifically, time in Switzerland is positively and highly significantly associated with both fluency in the local language (logistic regression, $p \ll 0.001$) and having a job (logistic regression, $p \ll 0.001$). In addition, we combined information about the participant's age and time in Switzerland to calculate and control for the proportion of the participant's lifetime spent in Switzerland. For example, an 18-year-old woman who moved to Switzerland when she was one has spent 17 years in Switzerland, while a 90-year-old man who moved to Switzerland as a five-year-old has spent 85 years in Switzerland. For both, the proportion of their lives in Switzerland is 0.944. Aside from our integration proxies, we also asked participants if they were from nomadic families. We did so in light of the hypothesis that families with a history in nomadism in Sudan would be favorable toward cutting because their nomadic lifestyles created a greater demand for the increased paternity certainty associated with infibulation (Mackie, 2000; Vogt et al., 2016).

Table 1 shows the results of regression analyses that model implicit association scores in Switzerland as a function of our control variables. Results show that the signs of estimated effects were stable for most variables regardless of model specification. Under the full model with the complete set of controls, two basic effects were significant. Specifically, older participants were significantly more positive about uncut girls. All else equal, an additional year of age was associated with an increase in D scores equivalent to 6.3% of a standard deviation. This result is consistent with our previous findings from Sudan, where older participants also held relatively positive attitudes about uncut girls (Vogt et al., 2016).

Time in Switzerland was also significantly associated with implicit attitudes, but in two different ways. First, the number of months a

Table 1

OLS regressions with implicit attitudes towards cutting (*D* scores) from Switzerland, normalized to have a mean of zero and a standard deviation of one, as a response variable. Standard errors, either robust (RSE) or not (SE), are shown in parentheses. Regressors include the age of the participant, the sex, variables indicating if the participant has at least one daughter and at least one son, months in Switzerland (CH), the proportion of one's life in Switzerland (CH), an indicator variable for German fluency, an indicator variable for being employed, and an indicator variable for whether the participant comes from a family of nomads. Breusch–Pagan tests for heteroskedasticity. Because the test does not detect significant heteroskedasticity when our cultural integration variables are included, we show results from the full model without robust standard errors in the last column.

Parameter	Estimate (RSE)	Estimate (RSE)	Estimate (RSE)	Estimate (RSE)	Estimate (SE)
Intercept	−0.183 (0.618)	−0.091 (0.215)	0.063 (0.109)	−1.802* (0.718)	−1.802* (0.685)
Age	0.0039 (0.021)			0.063* (0.025)	0.063* (0.025)
Female	0.297 (0.323)			0.295 (0.331)	0.295 (0.274)
Daughter(s)	0.249 (0.370)			0.209 (0.337)	0.209 (0.285)
Son(s)	−0.232 (0.264)			−0.177 (0.268)	−0.177 (0.283)
Months in CH		−0.0015 (0.0039)		−0.011* (0.0054)	−0.011** (0.0037)
Proportion life in CH		0.00099 (0.0069)		0.019 (0.011)	0.019* (0.0086)
Fluent in local lang		0.157 (0.359)		−0.034 (0.341)	−0.034 (0.275)
Employed		0.353 (0.384)		0.403 (0.327)	0.403 (0.255)
Family nomads			−0.353 (0.473)	−0.397 (0.512)	−0.397 (0.337)
Breusch–Pagan d.f.	9.800 4 p=0.044	6.358 4 p=0.174	3.887 1 p=0.049	11.951 9 p=0.216	

*** (0.001).
* (0.05).
** (0.01).

participant had lived in Switzerland was negatively associated with attitudes toward uncut girls. Second, the proportion of a participant's life in Switzerland was positively associated with attitudes toward uncut girls. The negative effect was significant regardless of whether we used robust standard errors, which are consistent to heteroskedastic errors (White, 1980), while the latter effect was only significant if we did not use robust standard errors. An ex post analysis of residuals indicates that robust standard errors may be overly conservative (Table 1, Breusch-Pagan, p=0.216), and so we treat both of these results as significant. Crucially, however, the negative result clearly dominates. In particular, all else equal, an additional month in Switzerland was associated with a decrease in D scores equivalent to 1.1% of a standard deviation. This translates into a decrease of more than 13% of a standard deviation per year in Switzerland. The positive effect associated with the proportion of one's life in Switzerland was small in comparison. Specifically, the difference between a proportion of zero, which means 0% one's life in Switzerland, and one, which means 100% of one's life in Switzerland, was 1.9% of a standard deviation. Even if we limit attention to the effect of just one additional year in Switzerland, the negative effect is significantly larger (Table 1, linear combination, $12\widehat{\beta}_{[\text{Months in CH}]} - \widehat{\beta}_{[\text{Proportion life in CH}]} = 0$, F=7.85, p=0.0065).

These results suggest that, if anything, time in Switzerland is associated with increasingly negative attitudes toward uncut girls. More broadly, the findings suggest that Sudanese immigrants in Switzerland are not more positive about uncut girls than people in Sudan because of a strong shift in attitudes after arriving in Switzerland. Rather, Sudanese immigrants in Switzerland could be a

special subsample of the Sudanese population (Farina & Ortensi, 2014; Ortensi et al., 2015).

The remaining estimated effects in Switzerland were not significant (Table 1), though in most cases they were broadly consistent with previous results from Sudan (Vogt et al., 2016). Specifically, in both Sudan and among Sudanese immigrants in Switzerland, we found that women and participants with daughters were relatively positive about uncut girls, while participants with a family history of nomadism were relatively negative about uncut girls. These results were significant with data from Sudan (Vogt et al., 2016), but not with data from Switzerland. Null results from Switzerland could obtain, of course, because a participant's sex, having daughters, and a history of nomadism are in fact unrelated to implicit attitudes in Switzerland. Alternatively, however, a potential concern with these null results is that our sample is small, and power was limited for these variables. Previous analyses of implicit association scores from Sudan (Vogt et al., 2016) indicate that a participant's sex, having daughters, and a family history of nomadism increased R^2 values by 0.02 or less. If we carry this value over to our Swiss context, a simple power analysis indicates that indeed we would have needed from several hundred to a few thousand participants to achieve a power of 0.8 with respect to a participant's sex, having daughters, and a history of nomadism. In Sudan we had nearly 8000 participants, but this is not possible in Switzerland. Available data indicate that considerably less than 1000 Sudanese live in Switzerland, many of them children. For this reason, the null results associated with a participant's sex, family composition, and family history of nomadism should *not* be interpreted as clear evidence of no effect.

5.3. Accounting for migration and acculturation

Our data from Sudan evinced a highly significant correlation between average implicit attitudes by community and the community-level risk of cutting for young girls (Fig. 4). In effect, average attitudes in communities are related to the risks of being cut in those same communities. This suggests that, if we view immigrants in Switzerland as a community, then extrapolating the cutting risk from Sudan to Switzerland without accounting for the difference in attitudes between them is likely an overestimate of the cutting risk in Switzerland.

Exterkate (2013) and Ortensi et al. (2015) account for such subtleties when estimating risk via the migration and acculturation impact factor. This factor captures the joint effect of selective migration out of the country of origin and acculturation to new norms and attitudes after immigration into a non-cutting country. Our results suggest that the migration effect is likely to dominate for Sudanese immigrants in Switzerland. In any case, whatever the balance of selective migration and acculturation, researchers have used the impact factor to place lower and upper bounds on the estimated risk of cutting among immigrants (Exterkate, 2013; Leye et al., 2014; Ziyada et al., 2016).

The upper bound on risk follows from assuming that selective migration and acculturation have no effect. Immigrants are assumed to be exactly like non-migrants in the country of origin. In terms of their implicit attitudes toward uncut girls, we have shown that Sudanese immigrants in Switzerland are not exactly like non-migrants in Sudan. A key question concerns what this attitudinal difference tells us about the risk of cutting in Switzerland. The attitudinal difference implies that the upper bound on risk in Switzerland is lower than the upper bound derived by assuming Sudanese people are the same in Switzerland and Sudan. This inference, however, requires the assumption that relatively positive attitudes toward uncut girls in Switzerland have translated into reduced risk.

Our Sudanese data show that across communities average implicit attitudes among adults are in fact related to the risks of cutting for girls. For this reason, if we compared two communities in Sudan, we would expect the community with relatively positive attitudes toward uncut girls to impose a relatively low risk of cutting. Although we do not know with certainty, we expect this effect on risk to be at least as strong when comparing Sudan to Switzerland. Apart from any pure effect associated with the attitudinal difference, all extrinsic forces are aligned against cutting in Switzerland. Local norms, the threat of legal sanctions, social pressure, and practical difficulties should reduce, all else equal, the probability of cutting in Switzerland compared to Sudan. When coupled with the observed difference in attitudes, our results imply that the upper bound on risk is very likely to be lower than what follows from assuming Sudanese immigrants are just like non-migrants in Sudan. This conclusion is broadly consistent with what Exterkate (2013) and Ortensi et al. (2015) concluded when estimating risk among immigrant populations in the Netherlands and Italy.

6. Discussion

Globalization and immigration place thousands of girls and women from countries with a long history of female genital cutting in countries with little or no history of the practice. In some of these latter countries, cutting is even illegal. We currently have little reliable information about attitudes towards cutting and the risk of cutting among immigrant populations in Europe and North America. The most prominent approach to estimating prevalence and the risk of being cut among immigrant populations is to extrapolate from countries of origin. If, however, immigrants have different attitudes from those of their counterparts in the country of origin, this kind of extrapolation is too simple. Using Sudan and Switzerland as an example, we have provided empirical evidence of exactly this kind of discrepancy. Sudanese immigrants in Switzerland have significantly more positive attitudes towards uncut girls than Sudanese people in Sudan. Moreover, our results suggest that Sudanese immigrants are relatively positive about uncut girls, at least in part, because of selective migration out of Sudan. Time in Switzerland is, if anything, associated with increasingly negative attitudes toward uncut girls. This leaves selective migration as a mechanism that probably contributes substantially to the relatively positive attitudes among Sudanese immigrants in Switzerland.

In particular, education tends to be associated with reduced support for female genital cutting (Exterkate, 2013; UNICEF, 2013). More highly educated mothers might be more likely to understand that cutting is not a religious obligation and can involve health risks. Interestingly, several studies have also shown that immigrants tend to be more highly educated than their home populations (Mafukidze, 2006; McKenzie & Rapoport, 2010). These two patterns suggest the possibility that relatively educated Sudanese could be both relatively positive about uncut girls and relatively likely to migrate to Europe. Though speculative, this combination would yield an outcome consistent with our data.

Aside from differences between countries, differences also exist within countries. Indeed, recent research shows that in many African countries attitudes related to female genital cutting vary substantially at the individual and household levels (Bellemare et al., 2015; Efferson et al., 2015; Hernlund & Shell-Duncan, 2007). Families who support cutting and families who do not essentially live in close proximity to each other. This tremendous local heterogeneity within countries of origin suggests an important question about immigrants. Namely, how are attitudes about cutting distributed within immigrant populations in Europe and North America? In addition to finding that Sudanese in Switzerland are more positive about uncut girls than in Sudan, we also found tremendous variation in attitudes among immigrants within Switzerland. The variation was significantly less than in Sudan, but still substantial, running the gamut from implicit attitudes that were strongly negative about uncut girls to those that were strongly positive. This kind of variation within an immigrant population has at least two important policy implications.

First, an immigrant's country of origin can provide, at best, a crude indication of any given immigrant's attitude toward cutting and the risk that her daughters will be cut. Uncritically casting suspicion on a group of immigrants is simplistic, and it could exaggerate misunderstanding and tension between immigrants and the people of European origin around them (European Institute for Gender Equality, 2016). Second, variation suggests tremendous potential for culturally sensitive programs that promote the abandonment of cutting among immigrants (Costello, Quinn, Tatchell, Jordan, & Neophytou, 2013; Merry, 2006). In particular, to the extent that a program ignores extant heterogeneity among immigrants, immigrants may feel prejudged, misrepresented, and attacked, all of which presumably increase the risk of backlash. Instead, programs could exploit the fact that various attitudes about cutting are already present within the immigrant population. In this sense the immigrant population provides a potential framework for the abandonment of cutting from within the population itself. Programs do not necessarily need to put forward facts, figures, and impassioned pleas that argue against female genital cutting from a perspective outside the immigrant culture. Our results suggest that contrasting viewpoints are already present among immigrants just as they are in countries of origin. Instead of indiscriminate suspicion and programs directed at entire immigrant communities, programs could be more effective if they actively build upon heterogeneity among immigrants themselves. For example, Vogt et al. (2016) recently showed that highlighting discordant views on female genital cutting within a cutting society can significantly improve attitudes towards uncut girls.

With respect to socially harmful practices like female genital cutting, our results highlight the importance of comparative studies directed at understanding the differences between migrants and non-migrants and the variation within immigrant populations. Policy makers need better data on the distribution of attitudes regarding female genital cutting in order to provide better estimates of cutting practices among immigrants and by extension better policy recommendations. Moreover, our study demonstrates how a robust measure of implicit attitudes can be used if programmers would like to evaluate interventions promoting the abandonment of cutting among immigrants in situ.

7. Conclusion

Sudanese immigrants in Switzerland have systematically more positive attitudes toward uncut girls than non-migrants, and selective migration out of Sudan likely contributes to this difference. Our study also found that migrants vary tremendously in their support for female genital cutting. In lieu of criticizing an immigrant culture from outside, the tremendous heterogeneity we observed offers an opportunity to develop culturally sensitive interventions that promote abandonment by turning to heterogeneous viewpoints on cutting already present within the cutting society. Not all immigrants have the same attitude toward cutting, and neither do they necessarily share the attitudes of people in their countries of origin.

Acknowledgements

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ssmph.2017.02.002.

References

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 to 10 and adulthood. *Psychological Science*, 17(1), 53–58.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful women: Does exposure reduce bias? *The Quarterly Journal of Economics*, 124(4), 1497–1540.

Bellemare, M. F., Novak, L., & Steinmetz, T. L. (2015). All in the family: Explaining the persistence of female genital cutting in West Africa. *Journal of Development Economics*, 116, 252–265.

Blattman, C., Jamison, J. C., Koroknay-Palicz, T., Rodrigues, K., & Sheridan, M. (2015). Measuring the measurement error: A method to qualitatively validate survey data. Technical report, *National Bureau of Economic Research*.

Boddy, J. (2007). *Civilizing women: British crusades in colonial Sudan*. Princeton: Princeton University Press.

Camilotti, G. (2015). 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.

Central Bureau of Statistics (CBS), UNICEF Sudan (2016). *Multiple Indicator Cluster Survey 2014 of Sudan*, Final Report.

Chauchard, S. (2013). Using mp3 players in surveys: The impact of a low-tech self-administration mode on reporting of sensitive attitudes. *Public Opinion Quarterly*, 77(S1), 220–231.

Costello, S., Quinn, M., Tatchell, A., Jordan, L., & Neophytou, K. (2013). In the best interests of the child: Preventing female genital cutting (fgc). *British Journal of Social Work*, 45(4), 1259–1276.

de Cao, E., & Lutz, C. (2015). Measuring attitudes regarding female genital mutilation through a list experiment. *Technical report*.

Dubourg, D., Richard, F., Leye, E., Ndame, S., Rommens, T., & Maes, S. (2011). Estimating the number of women with female genital mutilation in Belgium. *The European Journal of Contraception & Reproductive Health Care*, 16(4), 248–257.

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

European Institute for Gender Equality. (2013). Female genital mutilation in the European Union and Croatia. ⟨http://www.eige.europa.eu/sites/default/files/documents/eige-report-fgm-in-the-eu-and-croatia.pdf⟩.

European Institute for Gender Equality. (2016). Estimation of girls at risk of female genital mutilation in the European Union. ⟨http://eige.europa.eu/sites/default/files/documents/MH0215093ENN_Web.pdf⟩.

Exterkate, M. (2013). Female genital mutilation in the Netherlands: Prevalence, incidence and determinants. *Pharos*.

Farina, P., & Ortensi, L. E. (2014). The mother to daughter transmission of female genital cutting in emigration as evidenced by Italian survey data. *Genus*, 70, 2–3.

Gele, A. A., Johansen, E. B., & Sundby, J. (2012). When female circumcision comes to the West: Attitudes toward the practice among Somali immigrants in Oslo. *BMC Public Health*, 12(1), 697.

Gele, A. A., Sagbakken, M., & Kumar, B. (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*, 7, 933–943.

Glynn, J. R., Kayuni, N., Banda, E., Parrott, F., Floyd, S., Francis-Chizororo, M. et al. (2011). Assessing the validity of sexual behaviour reports in a whole population survey in rural Malawi. *PLoS One*, 6(7), e22840.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97(1), 17–41.

Gruenbaum, E. (1996). The cultural debate over female circumcision: The Sudanese are arguing this one out for themselves. *Medical Anthropology Quarterly*, 10(4), 455–475.

Gruenbaum, E. (2001). *The female circumcision controversy: An anthropological perspective*. Philadelphia: University of Pennsylvania Press.

Gruenbaum, E. (2005). Socio-cultural dynamics of female genital cutting: Research findings, gaps, and directions. *Culture, Health & Sexuality*, 7(5), 429–441.

Hänselmann, K., Börsch, C., Ikenberg, H., Strehlau, J., & Klug, S. (2011). Female genital mutilation in Germany. *Geburtshilfe und Frauenheilkunde*, 71(03), 205–208.

Hernlund, Y., & Shell-Duncan, B. (2007). Contingency, context, and change: Negotiating female genital cutting in the Gambia and Senegal. *Africa Today*, 53(4), 43–57.

Isman, E., Ekeus, C., & Bergen, V. (2013). Perceptions and experiences of female genital mutilation after immigration to Sweden: An exploratory study. *Sexual and Reproductive Healthcare*, 4, 93–98.

Jager, F., Schulze, S., & Hohlfeld, P. (2002). Female genital mutilation in Switzerland: A survey among gynaecologists. *Swiss Medical Weekly*, 132(19/20), 259–264.

Johnsdotter, S., Moussa, K., Carlbom, A., Aregai, R., & Essen, B. (2009). Never my daughters: A qualitative study regarding attitude change toward female genital cutting among Ethiopian and Eritrean families in Sweden. *Health Care for Women*

International, 30(2), 114−133.

Korfker, D. G., Reis, R., Rijnders, M. E., Meijer-van Asperen, S., Read, L., Sanjuan, M. et al. (2012). The lower prevalence of female genital mutilation in the Netherlands: A nationwide study in Dutch midwifery practices. *International Journal of Public Health*, 57(2), 413−420.

Leye, E., Mergaert, L., Arnaut, C., & Green, S. O. (2014). Towards a better estimation of prevalence of female genital mutilation in the European Union: Interpreting existing evidence in all EU member states. *Genus*, 70, 1.

Mackie, G. (2000). Female genital cutting: The beginning of the end, in: Shell-Duncan, B., & Hernlund, Y. (Eds.). Female circumcision in Africa: Culture, controversy, and change. Boulder, CO: Lynne Rienner Publishers, 253−283.

Mafukidze, J. (2006). A discussion of migration and migration patterns and flows in Africa. In Views on Migration in Sub-Saharan Africa. In: Cross, C., Gelderblom, D., Roux, N., & Mafukidze, J. (Eds.). Views on migration in Sub-Saharan Africa: Proceedings of an African migration alliance workshop. Cape Town: HSRC Press, 103-129.

McKenzie, D., & Rapoport, H. (2010). Self-selection patterns in Mexico−US migration: The role of migration networks. *The Review of Economics and Statistics*, 92(4), 811−821.

Merry, S. E. (2006). *Human rights and gender violence: Translating international law into local justice* Chicago: University of Chicago Press.

Morison, L., Dirir, A., Elmi, S., Warsame, J., & Dirir, S. (2014). How experiences and attitudes relating to female genital circumcision vary according to age on arrival in Britain: A study among young Somalis in London. *Ethnicity and Health*, 9(1), 75−100.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The implicit association test at age 7: A methodological and conceptual review, in: Bargh, J. A. (Ed.). Automatic processes in social thinking and behavior. London: Psychology Press, 265−292.

Ortensi, L. E., Farina, P., & Menonna, A. (2015). Improving estimates of the prevalence of female genital mutilation/cutting among migrants in western countries. *Demographic Research*, 32, 543−562.

Reig Alcaraz, M., Siles González, J., & Solano Ruiz, C. (2014). Attitudes towards female genital mutilation: An integrative review. *International Nursing Review*, 61(1), 25−34.

Rooth, D.-O. (2007). Implicit Discrimination in Hiring: Real World Evidence. *Technical report*, IZA Discussion Paper No. 2764.

Schroder, K. E., Carey, M. P., & Vanable, P. A. (2003). Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. *Annals of Behavioral Medicine*, 26(2), 104−123.

Shell-Duncan, B., & Hernlund, Y. (2000). 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. Boulder, CO: Lynne Rienner, 1−40.

Shell-Duncan, B., Wander, K., Hernlund, Y., & Moreau, A. (2013). Legislating change? Responses to criminalizing female genital cutting in Senegal. *Law & Society Review*, 47(4), 803−835.

Swiss Federal Statistical Office (2016, accessed 14 April 2016). Permanent and non permanent resident population by canton, sex, residence permit, age class and citizenship. ⟨http://www.bfs.admin.ch/bfs/portal/en/index/themen/01/22/lexi.html⟩.

Thomas, L. (2000). 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. Boulder, CO: Lynne Rienner, 129−150.

UNICEF (2013, accessed 16 September 2015). Female genital mutilation/cutting: A statistical overview and exploration of the dynamics of change. ⟨http://www.unicef.org/publications/index_69875.html⟩.

Vogt, S., Zaid, N. A. M., Ahmed, H. E. F., Fehr, E., & Efferson, C. (2016). Changing cultural attitudes towards female genital cutting. *Nature*. 538 (7626), 506-509.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 48(4), 817−838.

World Health Organization (2014, accessed 2 April 2016). Female genital mutilation. ⟨http://www.who.int/mediacentre/factsheets/fs241/en/⟩.

Yoder, P. S. & Khan, S. (2008). Numbers of women circumcised in Africa: The production of a total. *Calverton Maryland Macro International MEASURE DHS 2008 Mar.*

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

Ziyada, M. M., Norberg-Schulz, M., & Johansen, R. E. B. (2016). Estimating the magnitude of female genital mutilation/cutting in Norway: An extrapolation model. *BMC Public Health*, 16(1), 1.