

Cristina Bicchieri and Hugo Mercier

Norms and Beliefs: How Change Occurs

Societies are rife with negative, damaging practices, from open defecation to female genital cutting, endemic in many developing countries, to corruption and violence against women and children that we also witness in many Western societies. The theoretical and practical challenge we face is twofold. On the one hand, we want to explain what generates and supports such practices. On the other, we want to find ways to change them permanently. We will argue here that social norms play an important role in both tasks. Often norms support or embed certain practices, so that eliminating the latter involves changing the former. Sometimes, however, norms have to be created in order to eliminate a negative practice and support a new one, as we know of several widely practiced behaviors that are not supported by norms, but can be changed by introducing them. To understand what we mean by ‘practice’ and ‘norm’, we shall next refer to Bicchieri (2006) definition of social norms, a definition that allows to shed light on the way norms are supported, and on ways we may act to change them.

Social Norms

There are many behavioral regularities we engage in, from brushing teeth in the morning to adopting dress codes, from staying in line to buy a movie ticket to observing rules of fairness in allocating PhD slots. Some such regularities are behaviors that we adopt and keep following irrespective of what others do, or expect us to do. I brush my teeth every morning because I believe in certain hygiene principles, and the fact that most of the American population does the same has no impact on my decision. I care about germs and bacteria, not about what others do or don’t do. When I go to a party, however, I usually care about the local dress code, may ask about it, and try to conform to what I expect others to wear. Dressing differently would not

be a tragedy, just a cause for embarrassment, even if it is obvious that the other guests are tolerant and would not judge me negatively. In this case, what I expect others to wear has an influence on my decision about clothes. When I am in line to buy a movie ticket, I do not try to cut it or jump ahead. I expect everyone to patiently wait his turn, and I know that I am expected to behave accordingly. There is a sense that everyone *ought to* behave in an appropriate way, and we all get mad if someone tries to cut the line and jump ahead. Expecting this generalized reproach is enough to keep us all obeying the rule.

All of the above examples describe widely adopted behavioral regularities, the difference among them lying in the reasons why we follow them. In the case of dressing codes and staying in line, our expectations about what others do or will do are paramount in giving us a reason to behave in that particular way. Yet there is a difference between a simple empirical expectation (all will wear a black tie, all wait in line) and a normative expectation (all those who wait in line believe I ought to wait patiently in my place). In the first case, expecting a certain behavior gives me a definite reason to follow it; in the latter case, I need a further inducement in the form of a sanction (negative, in this example) to decide that it is better not to cut the line. Social norms, it has been argued (Bicchieri 2006), are behavioral rules supported by a combination of empirical and normative expectations. Individuals have a *conditional* preference for obeying social norms, provided they hold the right expectations.¹

Empirical expectations are always important, since in their absence we may be tempted to disobey social norms, especially those that demand behavior that may conflict with self-interest. Norms of cooperation, reciprocity, and fairness, for example, may lose their grip when we are faced with widespread transgressions. In that case, the force of the norm is greatly diminished. Yet, even when widely followed, social norms may require, to be obeyed, the further belief that others think we ought to obey them, and may be prepared to punish our transgression. Such *normative expectations* always accompany social norms, and are usually consistent with our empirical expectations of widespread compliance.

¹ Conditional preferences distinguish social from moral or religious norms, where one would choose to conform irrespective of what others are expected to do, or think one ought to do.

As we shall see, conceiving of norms as supported by, and in a sense constituted of, individuals' expectations offers many theoretical advantages. For one, we now have an operational definition of 'social norm' that allows us to make predictions and to experimentally test whether a change in expectations results in a change in behavior. We can also assess the presence of social norms by asking people about their second-order beliefs about what others think the appropriate behavior is, and check for the mutual consistency of these beliefs (Bicchieri & Chavez 2010). We can, and this is the topic of this paper, devise specific interventions to effect norm change by acting upon the expectations that support the norm we wish to eradicate or, when it is a new norm we want to establish, work at creating new expectations, and focus on those factors that will bridge expectations and behavior. If indeed expectations, both empirical and normative, are crucial to the existence and stability of a norm, it follows that a change in expectations will always induce a change in compliance and, when the change in expectations is widespread, the abandonment of a norm. For those interested in the removal of a negative norm, or the establishment of a new, positive norm, the issue of collective belief change thus becomes of paramount importance.

Changing Empirical Expectations: The Pitfalls

How easy is it to change people's empirical expectations? First of all, individuals should *observe* or at least reasonably *expect* different behavior in a large enough number of relevant people (i.e., people whose behavior and judgment they care about). Notice that there are many cases in which such observation/expectations would prove difficult to come by. Take for example norms about private behavior, such as sexual mores. In this case, we may have widespread, private disagreement with the standing norms, and a significant amount of secret deviance (Schank 1932). Yet, because public deviance may be costly, we would observe public, open allegiance and support for the norms in question. These cases are typical of *pluralistic ignorance*, a cognitive state in which one believes one's attitudes and preferences are different from those of similarly situated others, even if public behavior is identical (Allport 1924; Miller & McFarland 1991). In all these cases, individuals engage in social comparison with others who are similarly situated. Others' behavior is observable, or at least the consequences of behavior are observable, in that if there are few or no pregnancies out of wedlock one would be justified

in assuming that sex outside marriage is uncommon and condemned. In all these cases, transparent communication is impossible, as the social situation is one in which the norms in question are thought to be widely adopted and strongly endorsed, hence the fear of embarrassment and ostracism that would follow an open declaration of disagreement keeps people in line. Typically people assume that others' behavior is consistent with their attitudes and preferences, therefore from observing widespread compliance each will infer that everybody else endorses the social norm, which in turn can only reinforce public allegiance to it.

Such cases of pluralistic ignorance are quite common, even when behavior is public (as opposed to private), such as Prohibition support (Robinson 1932), the "conservative lag" in behavior toward integration (O'Gorman 1975), or a "liberal leap" such as the sexual revolution in the 1960s (Klassen et al. 1989). Studies of gang members (Matza 1964), prison guards (Klofas & Toch 1982) and prison inmates (Benaquisto & Freed 1996), as well as school teachers (Packard & Willower 1972) show that the social norms about proper behavior that are widely shared by all these communities are often regarded by their very members as too strict or even plainly wrong, but nobody dares to question the shared rules for fear of negative sanctions. It has been shown (Bicchieri & Fukui 1999) that it may take a small number of 'trendsetters' who question the standing norm and start behaving differently to effect a major change. But this would mean that we have to move our explanation a step up, in that we need to explain how change in behavior for the trendsetters came about.

Another possibility is change that comes, so to speak, from above. Imagine the case of a government injunction: From now on, female genital cutting (FGC) is abolished. We have plenty of experience, especially with developing countries, that such injunctions rarely work. It is interesting to note that, on the contrary, an injunction to shift driving to the right side of the road would (and has been) completely successful. Why? A widely announced change in traffic rules is expected to be followed by all drivers. It is in the interest of each individual driver to coordinate with others and knowing that, one can trust that other drivers will comply with the injunction to drive, say, on the right side of the road. This case is one of a shift in conventions. As discussed elsewhere (Bicchieri 2006), conventions are quite different from social norms, in that they are supported by empirical expectations of compliance, and a preference to follow the convention provided one expects

most (or all) others to comply with it. Thus a government diktat would work for conventions, but be more problematic in case of social norms.

To move away from a shared norm, we need the assurance that we will not suffer negative consequences. This is because social norms are also supported by normative expectations, i.e., the expectation that others believe we *ought* to behave in a given way, and may sanction us (in a negative way) if we stray. Changing norms thus presents us with a collective action problem, as nobody wants to incur the negative sanctions involved in a transgression. *Prima facie*, it would appear that external interventions, in the form of government interventions, may *facilitate* behavioral changes, by taking away the stigma connected with disobeying a widely held social norm. For example, if FGC is widely practiced in a community, then being the first to abandon it would expose the family to significant damages. For one, the uncut girl would not find a husband, and would become the target of negative perceptions.² The entire family would suffer negative consequences, as a family that does not cut its girls would be seen as openly flaunting shared norms, and would thus be ostracized.

It would thus seem that introducing laws that prohibit that practice, and thus establish new sanctions, would be a successful measure, as it would alter the cost and benefit of the targeted behavior by changing expectations and the perceptions of what incurs disapproval, and even change a person's own preferences and create guilt, especially when there is a shared norm of obeying the law. Public opposition to the existing norms would become less costly, and therefore we should see the target behavior eventually disappear. This view embodies the traditional economic analysis of law, an analysis that focuses on its role in changing the cost and benefit of targeted behavior: people are predicted to abide by the law if sanctions are sufficiently severe and tend to break the law if sanctions for doing so are too mild. Yet this view is too simplistic, in that it assumes a host of conditions that need to be present in order for the legal solution to be effective. The question whether laws bring about social change hinges on factors such as legitimacy, procedural fairness, and how the law is originated and enforced.

People who view the law as legitimate are more likely to comply with it even though this contradicts their interests. A legitimate law is not just

² The Saleema case study in Somalia points to the fact that the only word traditionally used for the uncut girl was “ghalfa,” which roughly means prostitute (Hadi 2006).

one that ensues from a legitimate, recognized authority. It must also be the case that the procedures through which authorities make decisions are seen as fair, that the law is consistently enforced, and that the enforcers are perceived as honest. So for example the sporadic campaigns that are launched to enforce the laws during politically sensitive periods, such as in pre-electoral times, are not taken too seriously, and the corruption of local enforcers is a powerful delegitimizing influence. Furthermore, individuals' opportunity to take part in the decision-making process, present their arguments, being listened, and having their views considered by the authorities would seem to offer an especially strong incentive to abide by the law.

Legitimacy thus results in respect for the authorities, and a sense of obligation to obey them. Yet, even assuming that the authority that enacts and enforces the law is perceived as legitimate, perhaps the most important factor that determines successful enforcement is a shared sense that the existing legal arrangements are *as they ought to be*, in that they do not appear so distant from existing social norms as to lose credibility.

If the law strays too far from the norms, the public will not respect the law, and hence will not stigmatize those who violate it. Loss of stigma means loss of the most important deterrent the criminal justice system has. (Stuntz 2000, 1872)

In other words, the law should approximate popular views, otherwise the threat to seek enforcement will not be credible. Platteau (2000) and Aldashev et al. (2010a,b) give a series of examples of laws that were successful precisely because they were sufficiently close to shared social norms: in Gabon and Senegal, instead of banning polygamy, the initial marriage contract allowed the choice of monogamy or polygamy. In Ghana, to protect women and children's inheritance rights, a moderate law proved more effective than previous extreme law. In Bogota, where high firearm mortality was common, Mockus, the mayor of Bogota, decided to ban guns on weekends only, sending a strong signal but also realistically understanding that a moderate legal injunction would be easier to enforce and obey. Dan Kahan's (2000) discussion of 'gentle nudges' and 'hard shoves' similarly points out that if a new legal norm imposes harsh penalties against a widely accepted social norm, police will become less likely to enforce the law, prosecutors will be less likely to charge and juries to convict, with the effect of reinforcing the existing norm that we wanted to change. Milder penalties are much more

effective, and enforceable, thus leading to a progressive condemnation and abandonment of the ‘sticky norm’.

In sum, the legal approach to norm change can help change empirical expectations, but only under rather strict conditions. Individuals will abandon a shared social norm only if they believe that others are changing, too. This belief must be accompanied by a credible change in sanctions, in that the original negative social sanction for *not* following the norm will be substituted by a new, credible negative sanction for following it. In this case, normative expectations would change, too.

Deliberation

A stepping-stone in the process of norm change is affecting people’s empirical expectations. If someone believes that others will act in a certain way with regard to a given norm – follow it, say – that person is likely to follow it herself, irrespective of whether she thinks this is the best thing to do otherwise (Bicchieri & Muldoon 2010). Prima facie, it may thus seem that what really matters is behavior: what people do, not what they say ought to be done. After all, as economists are wont to point out, talk is cheap (Farrell & Rabin 1996). Yet numerous experiments have demonstrated the power of discussion to promote pro-social behavior by focusing participants on ‘good’ norms (see, for review, Balliet 2010; Sally 1995).

One of the contexts in which norms play an important role is commons dilemmas. In her pioneering study of the emergence of social norms, Edna Ullmann-Margalit (1977) observed that norms help solve prisoners’ dilemma-type situations (and commons dilemmas are a type of PD). Indeed, she claimed that what she dubbed *PD-norms* emerge precisely to stifle the conflict between individual and common interests. Yet the question remains what individuals will do in situations in which a norm of cooperation exists and could be followed, but there is no sanction to keep them in line. Experiments on commons dilemmas are indicative of the difficulty of achieving and maintaining cooperation, but also suggest that a solution to the problem of norm compliance exists.

In an idealized, laboratory version of a commons dilemma, participants are given some endowment money and a choice either to keep that money to themselves or invest it in a common pool. All the money invested in the common pool is then multiplied by some amount (larger than 1) and

then equally redistributed across all participants. Overall profits are maximized when everybody contributes their whole endowment. Yet each individual is better off letting all the others contribute while keeping her endowment to herself. As a result, when the game is played in repeated rounds, contributions to the common pool rapidly decrease to a negligible level. When participants are not allowed to talk to each other, that is. If the participants can communicate prior to making their decisions, the level of cooperation can remain very high for as many rounds as the experimenter is willing to go (Ostrom, Walker, & Gardner 1992). When participants can communicate, they are able to focus on (and follow) a norm of contribution to the common pool.

What happens in these discussing groups that make participants more likely to cooperate? Part of it is the result of low-level factors; simply interacting with other people from the group makes participants more likely to cooperate with them, even if that interaction is as minimal as looking each other in the eyes (Kurzban 2001). But several experiments have demonstrated that the bulk of the effect comes from the ability to make promises (e.g., Orbell, Van de Kragt, & Dawes 1988; Bicchieri 2002). During the discussion, people promise to contribute a given amount and the evidence suggests that the majority of the participants are true to their word. In terms of norms, the effect of promises can be described as a change in empirical and normative expectations. Participants now expect others to behave in a way consistent with their pledges, and they expect that people who renege on their promise will be negatively judged.

A limitation of these experiments is that participants do not have to fight an ingrained norm that would hinder the acceptance of a norm of cooperation. What if there is a pre-existing norm that dictates no (or low) cooperation? This is not as farfetched as it may seem. In some cultures, people who contribute too much to a common pool are seen as exerting an undue pressure on others to match their level of contribution and are punished for it (Herrmann, Thöni, & Gächter 2008). In these circumstances, promises may not be sufficient, for they would be less credible. Knowing that sanctions can be incurred if the promises are kept, other participants may not take them as seriously as they would otherwise. As a result, their empirical expectations may remain unchanged, and they would not be inclined to follow a norm of cooperation.

If we look at real life cases, the risk of empty promises is even more blatant. Female genital cutting (FGC), mentioned above, provides a good example

(LeJeune & Mackie 2009). It consists in the ablation of parts of the female genitalia (usually the clitoris, sometimes more) and is typically practiced on relatively young girls, certainly before they get married. The practice of FGC is not an isolated cultural norm. It is embedded in a rich network of beliefs – beliefs about the origins of FGC, its religious justifications, its effect on health (or lack thereof) and so on. Many of these beliefs are normative in nature. The virtue of uncut girls, in particular, is often questioned. As a result, people follow the norm not only because of their empirical expectations – they expect others to do the same – but also because of their normative expectations – they expect to suffer from a variety of sanctions if they fail to follow the norm. In such a context, promises are much less likely to result in a switch in empirical expectations, as they are not very credible. The whole network of beliefs surrounding FGC – responsible for the normative expectations – cannot simply vanish; and, as long as it is present, it is going to make norm change extremely difficult. Even if empirical expectations are a crucial element in norm change, changing empirical expectations without first modifying normative expectations is not always possible.

Discussions and deliberations can also play a critical role in changing normative expectations. The simplest type of change that discussions can bring about is lifting pluralistic ignorance. As mentioned above, people can follow a norm because they believe that others would shun them if they didn't, even if this belief is mistaken. If people were only able to candidly share their feelings about the norm, they may just realize that the whole thing is pointless and stop abiding by it. The solution could therefore be purely endogenous. Often things are unlikely to be that simple though – if a friendly chat would have solved the problem, it is likely that the despised norm would have already disappeared. There are several reasons why the relevant exchange does not take place. In contrived laboratory situations – but also in a few real life cases – communication may simply be impossible. But the most common hindrance to a candid discussion of the norm is the existence of norms that dictate how one should talk about norms. Going back to the FGC example, even if we assumed that a sizeable part of the population was in fact opposed to the norm, these people would have very little chances of expressing such a view. This could be either because a specific segment of the population – women, often – is not given much of a public voice, or more drastically because the mere mention of FGC would be a very serious normative breach (LeJeune & Mackie 2009).

When the norm cannot be freely discussed by all the parties involved, trying to force people to talk about it anyway is likely to backfire. The external agent trying to impose such a discussion would likely be perceived very negatively. Even if the discussion were to take place, it could have damaging consequences. If criticisms of the norm are not allowed, a false impression of consensus can strengthen pluralistic ignorance. Following a discussion that all parties believe to have been frank – except for their own contribution – the norm could even acquire more legitimacy. An exogenous element is thus often required to challenge normative beliefs, either to challenge the normative beliefs themselves, or at least to question the normative beliefs that regulate how the targeted normative beliefs are discussed.

The role of the exogenous agent will be, simply put, to make people change their mind about the relevant normative beliefs. One way to do so is to rely on trust and authority. If a religious or secular leader tells people that some of their normative beliefs are mistaken, they may just take her word for it – especially if the leader is respected by everyone in the relevant community. But in many cases it is not possible to merely rely on trust: people have to be *convinced* that they should change their mind. The main tool for conviction is argumentation, and we presently give a brief account of how arguments can change people's beliefs.

Arguments and Belief Change

Beliefs rarely come by as isolated units; they form complex networks with different types of relationships: consequence, explanation, association, etc. It is possible to describe many of these links in terms of coherence: beliefs are more or less coherent, or consistent, with each other (Thagard 2002). Inconsistencies are typically the occasion for belief change. When inconsistent beliefs are detected, the mind tries to determine which can be most easily rejected in order to reduce the inconsistency (Festinger 1957). People can stumble upon these inconsistent beliefs on their own, or they can be made to face their inconsistencies by others. This is what arguments do. Arguments take a belief that the listener accepts – the premise – and show her that this belief is inconsistent with the rejection of the argument's conclusion. When a good argument is offered, it is more consistent for the listener to change her mind about the conclusion than to accept the premise while rejecting the conclusion (Mercier & Sperber 2011).

Arguments can be more or less explicit. In a very explicit argument, the logical relationship is highlighted with logical connectives (“and”) or other connectives (“therefore”). That the strength of the argument should be prominently on display is generally a good thing: it makes the argument easier to understand and more persuasive. Yet explicit arguments can also backfire. If the intent of the speaker is ambiguous in the first place, it is more likely to be perceived as manipulative (Kamalski et al. 2008). Moreover, explicit arguments may appear threatening. The listener may be unable to muster a sound counterargument while still not being persuaded. Such a situation is likely to arise when the issue is heavily emotionally loaded, as in cases of “moral dumbfounding” (Haidt, Bjorklund, & Murphy 2000). The listener is then likely to feel threatened by the argument, and to have an antagonistic reaction to the speaker who is challenging her beliefs and making her look irrational.

Arguments can also be mostly implicit. Instead of explicitly making the speaker face her inconsistencies, she can be led to realize on her own that some of her beliefs are in fact conflicting with each other. Social norms are steeped in a thick network of beliefs, attitudes, and values. Some of them are more central than others, and highlighting conflicts between beliefs (as well as between beliefs, values, and attitudes) must be threaded lightly. Tostan is a non-governmental organization that helped abandonment of FGC in Senegal, and they rely in large part on this type of implicit arguments (Gillespie & Melching 2010). They do not bluntly tell people that their beliefs about FGC are inconsistent with, for instance, their desire to have healthy children. Instead, the importance of some values – such as respect for human life – is first highlighted without reference to FGC. People are made to work out, in a process of collective deliberation, the practical consequences of these values. When this approach is coupled with information about FGC – in particular its health consequences – people can work out on their own the problematic aspects of FGC (Diop et al. 2004)

One of the factors that make some beliefs – such as beliefs about FGC – difficult to change is that they are more “central” than others (see, e.g., Judd & Krosnick 1982). These beliefs are at the center of a dense network of beliefs, attitudes, and values. Keeping on with the example of FGC, the belief that girls should be cut has explanations and consequences, it may be linked with religious beliefs and social customs, it is embedded in specific rituals, etc. A frontal attack on FGC is unlikely to succeed, as many other

beliefs would have to simultaneously evolve. By contrast, more peripheral beliefs are more amenable to arguments. For instance, the belief that FGC is part of the Islamic faith is peripheral both to beliefs about FGC and about Islam (this belief is a rationalization, as Islamic scriptures do not in fact recommend FGC). One of the reasons that tackling a relatively central belief often entails a prolonged process is that many peripheral beliefs have to be modified first. A complementary way to target relatively central beliefs is to use beliefs that are even more central. This is one way of describing a major aspect of Tostan's work with deliberations: trying to show that some central values conflict with common beliefs, such as the belief that girls should be cut.

Discussions and deliberations often allow people to change their normative beliefs and, therefore, the normative expectations related to an old norm that has to be challenged. Still, even the disappearance of the previous normative expectations may not prove sufficient. There are several reasons this may occur. Agreement that a particular norm is not necessary to fulfill some core beliefs, and indeed may be in conflict with some deeply held values is just a first step, necessary but by no means sufficient, to change behavior. People must be convinced that their core beliefs and values are better served by a new practice. Such new practices may be endorsed by a respected leader, or be the result of extensive group discussion that focuses on alternative solutions. The importance of finding alternatives cannot be overstated. Without the possibility of conceiving viable alternatives, abandoning an established norm is a losing proposition.

In lengthy rounds of collective discussions, people may agree that the old norm should not be upheld, come to envision and agree upon a new practice, and promise to follow it (Haile 2006). Yet if the consequences of being isolated in keeping one's word are too high, people may be reluctant to do so – especially since they know that others are likely to have the same train of thought and therefore to back down as well (a reflection that can be made worse by iteration). In communities that practice FGC, it is virtually impossible for an uncut girl to find a husband. In other words, the costs of not following the old norm are potentially enormous. Even if everybody can be persuaded to promise to forswear the custom, people may still fear that others won't keep their word. What is needed then is the establishment of normative beliefs that will transform the new, agreed upon practice into a social norm.

One way to enact such transformation is to publicly commit to change behavior and promise to move in the newly envisaged direction. Public pledges have many advantages: the promisor is more likely to keep his or her word since not doing it exposes to ‘loss of face’ and possibly also to reputational damage. Knowing the costs of a broken promise makes it credible, and generates the trust necessary to start moving in the new direction. Furthermore, even if some participants are not particularly enthusiastic about the new course, witnessing a large number of people committing to change behavior leads to form new empirical and normative expectations. Public, credible promises have the function of creating a common belief that the new behavior will be implemented, and the *expectation* of such behavior. Creating normative expectations, however, is crucial in establishing the new behavior as a social norm.

In an experiment alluded to earlier, participants were able to reach a high level of cooperation – high contributions to a common pool – simply by discussing the game among themselves and making promises. However, a simple variant of that experiment reveals the limits of simple promises. When the stakes were higher, making defecting more appealing, promises were much less successful at maintaining cooperation (Ostrom et al. 1992). A possible solution is to develop normative expectations by introducing sanctions against defectors. Indeed, participants are willing – eager even – to inflict punishment on defectors, in spite of personal costs. One of the reasons punishments are effective in simple commons dilemmas is that their meaning is usually unambiguous. If a participant breaks her word to contribute at least a certain amount, she is likely to understand why she is then punished. Most real life situations, however, are more intricate, so that one could be punished without knowing exactly why. When this happens, the individual being punished may not know how to improve her behavior or, even if she understands why others think she should be punished, she may be unwilling to change, as she may perceive the punishment as unfair.

Discussions and deliberations can also play a crucial role for the establishment of ways to enforce commitments. When a group of people is trying to institute a norm, they are likely to realize that some sanctions for norm-breakers are in order. Deliberation is a good way to devise monitoring and punishment devices. If the group members have different incentives and perspectives, their views can be heard and taken into account. The resulting sanctioning scheme will be perceived as more legitimate, and will therefore

be more effective. Discussions can also prove critical when the punishment is inflicted, as they facilitate an understanding of why it is inflicted and how it can be avoided in the future (Janssen et al. 2010).

Common Knowledge and Tipping Points

For most beliefs the most effective way to change them, and thus eventually change the practices that they support, is through argumentation. For argumentation to be successful, however, two conditions are required. First, the arguer must be able to rest on a set of explicit beliefs and values that is equally well entrenched in the listener and that is inconsistent with the target belief that we want to change. Second, the belief must not be held mainly because other people hold it as well. In this latter case, argumentation is not likely to succeed: as long as one does not see other people from the relevant group changing their mind, one is unlikely to change her beliefs. Social norms, we have argued, are supported by *shared* normative beliefs. Therefore the process of belief change has to be a *collective* one. People, in other words, have to change their mind together. Group discussion, as opposed to individual discussion, is important because, if a group is confronted with a persuasive argument, and people see others accept it, then they may feel free to accept the argument themselves. Accepting an argument and changing behavior, however, are two different things. One may be convinced by an argument and change one's attitude towards a given norm, but hesitate to change behavior for fear of being in a minority. This means that trying to change the behavior of one person after another is bound to be extremely difficult, if not impossible. For a norm to change, the whole group – or at least a sizeable majority – must be reached.

Deliberation and group diffusion are two complementary and necessary ways to make change happen. Yet there is a tension between the two. We know that deliberation works best in small group settings, but if the relevant group is large, using the 'common knowledge of change' approach requires the entire group to change its mind. In the successful Tostan experience, deliberation in small core groups reaches conclusions that are unstable unless and until the others follow. Members of the small group have an incentive to recruit more people up to the point at which enough people are ready to adopt a new practice. Typically, the core group organizes diffusion of their discussions into wider arenas. In the African experience so well exemplified

by Tostan, diffusion has taken several forms: ordinary discussions with family and friends; meetings with elders, religious leaders, and the women's group; a meeting of the whole community; discussions in nearby communities; and inter-village meetings with delegates from surrounding communities. Spontaneous diffusion, when we let the information circulate of its own accord, often cannot be relied upon until the last phase of the operation.

When a practice is strongly interdependent, often it is not enough for individuals simply to adopt a more favorable *attitude* towards a new alternative. The greater the loss (for example, damage to the daughter's reputation) resulting from a failed effort to shift to a new norm, the more people need to be sure that a sufficient number of people in the community will together *act* to adopt the new alternative. All must see that all see that there is change. Since norms are grounded on expectations, what we think others do, and what we think others think we should do, must both change in order for a fracture with the past to occur. We engage in alternative behaviors only if we think other people do so as well, and will judge us well for it. Within a population, it often happens that not everybody follows a norm. When this is the case, people can take the proportion of the population that follows a norm into account in their decisions. Imagine a population in which most people would prefer not to beat their wives, but there's a tradition – a norm in fact – to beat them for even small misdeeds. Furthermore, ideals of masculinity, honor, and family values are deeply linked to the practice. At some point, a few individuals may be convinced that beating wives is not the best way to fulfill deeply held values, and they may even decide to abandon the practice. Most others remain unmoved, as the minority is too small. Here core group discussion and organized diffusion would play a crucial role, effecting a gradual change in attitudes. If the minority keeps growing, it may reach a tipping point. At this stage, the minority has grown large enough that most other people feel free to break from the norm and stop beating their wives. Norms often change in this way. Progress is very slow at first, as a few people gradually start to adopt a new norm. But when the tipping point is reached, change can be very sudden. It should then be expected that a slow and steady change in attitudes may not be immediately accompanied by an equally slow and steady change in behavior. On the contrary, behavioral change may be sudden and quite dramatic, and difficult to predict. In the experiences that have accompanied abandonment of FGC, change typically occurs when the population reaches common knowledge that a majority is

ready to abandon the old practice. Everybody knows that everybody else knows that the majority of the population is adopting a new practice. There are many ways in which such common knowledge can be achieved: an elaborate public declaration by representatives of interconnected communities; the posting and propagation of a decision by a respected and effective local political authority, or the signing of a flag symbolizing the change by each household in the community. All these are ways to publicly celebrate the change and let everyone know that new expectations are in place.

From the Lab to the Field: Scaling up Norm Change

In our analysis of norm change and deliberation, we have relied substantially on laboratory experiments, accompanied by real life examples. Following the lead of scholars such as Elinor Ostrom (1991), we urge for a better integration of fieldwork and laboratory experiments. The results obtained in the laboratory, often with so-called WEIRD participants (participants from Western Educated Industrialized Rich Democratic countries, see Henrich, Heine, & Norenzayan 2010) can not always be generalized to the field. While it is easy to conjure up examples of disappointing group performance in real life – dreadful committee experiences are burned in our memories – we would like to provide an example in which groups in the field can avoid the pitfalls in which their laboratory counterpart regularly fall.

When psychologists ask participants who agree on some issue to talk about it anyway, the attitudes of the participants tend to polarize. For instance, a group of Republicans talking about a tax increase are likely to pile up arguments against it, arguments that everyone is likely to accept uncritically, providing group members with even more reasons to reject the tax hike. Group polarization is so reliably observed in the lab that Cass Sunstein saw it fit to turn it into the ‘*law of group polarization*’ (2002). Yet the analysis of real life cases fails to back up such a strong generalization. Historical analyses of important decisions have shown that groups can start and continue being very cohesive, sharing in the same ideology, without succumbing to groupthink and the polarization that generally ensues (Tetlock et al. 1992). Studies of ‘enclave’ deliberation among disempowered groups have shown that despite a lack of heterogeneity, deliberation can allow such groups to find a voice without leading to groupthink or polarization (Karpowitz, Raphael, & Hammond 2009). A prominent historical example

is that of the close-knit group of Quakers who, despite their agreement on the fundamental issue, did not polarize, putting forward instead pragmatic solutions that helped achieve the abolition of slavery in England (see, e.g., Brown 2006). In-depth studies of such cases are necessary to understand in what respect they differ from the laboratory situations that so reliably produce polarization. One suggestion may be that when the personal stakes of the group members increase, polarization is less likely to ensue. Such a hypothesis would greatly diminish the relevance of the laboratory results obtained so far, but not of experiments in general. Indeed, it would be necessary to test the hypothesis in the laboratory to establish its validity.

As the example of group polarization shows, one must exert caution when extrapolating from the laboratory to the real world. A better interaction and integration of field studies and laboratory experiments will be necessary if we are to reach conclusions that are both sound and relevant.

There is, however, another drawback that the studies we have cited so far share, whether they have been done in the lab or in the field: their relatively small scale. Laboratory experiments only involve a very limited number of participants at a time. The example we have taken of the role of deliberation in norm change – the work of Tostan in supporting the abandonment of FGC – involves significantly larger groups – up to 150 people or more. Yet even with groups of that size, it is relatively easy to imagine that deliberations can affect a substantial part of the group either directly or with at most one level of communication (i.e., someone involved in the debate talking about it with someone who hadn't taken part in it). Given that norm change requires that a substantial section of the population is ready to change – the famed tipping point – it is not clear exactly how such a process can be scaled up if the goal is to change a norm in an average Western country of several millions inhabitants. This issue is particularly important for policy makers who want to promote behavioral changes in areas such as health or business where entrenched norms stand in the way of progress.

Some processes of norm change are susceptible to scale up relatively easily. Imagine for instance a typical situation of pluralistic ignorance. If people can be made to speak their mind more or less freely when they are surveyed by a pollster, and if the results of the poll can be made public by a trusted source, there is hardly any limit to the size of the population that can be affected. The only constraints are the costs of polling and publicizing the results. In many cases, however, the change has to be deeper than simply

making people reveal their true preferences: people have to genuinely change their mind. Deliberation is the best tool to induce belief change, but it doesn't scale up very well: Studies show that as groups grow larger, pre-play communication aimed at inducing cooperative behavior breaks down more easily, as it becomes more difficult to create the trust necessary to support commitments to cooperate. Large groups, it has been argued, could benefit from computer-mediated communication. Yet, even with small numbers, we know that cooperation is more difficult to establish when the means of communication is a computer (Bicchieri & Lev-On 2007; Bicchieri, Lev-On, & Chavez 2010). Important aspects of 'commitment production', such as coordinating mutual promises, the credibility of promises, and attainment of public knowledge about mutual promising, become problematic in computer-mediated environments. If mutual expectations are crucial in attaining belief (and norm) change, finding the means to achieve a change in expectations should be one of our main goals.

The difficulties we have highlighted mean that deliberation can hardly be the only means through which a norm can change when large numbers of people are involved. Activists understand this very well, and so they rely on a variety of other media to effect norm change, from ad campaigns to spreading new words that encapsulate a normative statement (such as "homophobic").

Deliberation, however, should not be written off when large numbers are involved. One of the most important movements in recent political science pushes for a more "deliberative democracy" (see, e.g., Elster 1998; Gutmann & Thompson 1996). Partisans of deliberative democracy are obviously aware of the scaling up problem, but they can be willing to confront it head on. For instance, Ackerman and Fishkin have suggested that a national "Deliberation Day" should be instituted (Ackerman & Fishkin 2004). During this day, which would be a national holiday held shortly before an election, all the registered voters would be invited to discuss their views on the upcoming election. While such a project may sound unrealistic now, smaller versions of the same idea have already been implemented. For instance, *AmericaSpeaks* organizes debates between small groups of citizens, who then share their results with a larger local group of several hundred people, who then shares these results with other such groups across the country, reaching several thousand people. The goal of such deliberation is not to effect norm change directly, but they can be a crucial step on the way to norm change. People

can get a better idea not only of what other people think, but also of why they hold such views (Hansen 2003). More importantly maybe, people can change their mind about norm-relevant beliefs (e.g. Luskin, Fishkin, & Jowell 2002).

Aside from these formal debates, deliberation can also play a critical role in norm change through its action in everyday life. Mansbridge has argued that students of deliberative democracy should pay greater attention to the role of “everyday talk” (Mansbridge 1999). In her study of the feminist movement, she has noted how women have been able to exert an influence on men in their surroundings through ordinary interactions. Such local interactions are influenced by larger trends. For instance, women were able to recruit terms devised by activists – such as “male chauvinist” – in order to make a point quickly and effectively. But, importantly, the multiplication of similar local interactions can also exert a significant effect on the population at large.

Clearly, a lot remains to be done to link the study of local interactions with the application to norm change in large polities. We hope that a better integration of lab experiments and field data, as well as increased dialogue between psychology and the social sciences will help close that gap.

Conclusion

To abandon negative norms, we need to change people’s empirical and normative expectations. Discussions and deliberations can be effective means to enact change, as they facilitate the creation of the new empirical and normative expectations that are central to a norm’s existence. The positive side-effects of collective deliberations, such as improved interpersonal understanding (Fishkin & Luskin 2005), increased respect among participants (Gutmann & Thompson 1996), better solutions to a variety of practical, moral, and intellectual problems (Mercier 2011; Mercier & Landemore 2012) can prove significant to norm change, but are likely to be peripheral to the main issue. In the present paper we have confined our analysis to aspects of discussions and deliberations that allow tackling norm change more directly, starting with the ability to change people’s empirical expectations. Discussions can change attitudes and clarify what people intend to do. Norm change can sometimes be effected simply by people promising that they will abandon the old norm or follow a new one. For promises to be effective,

however, they have to be credible. If people think that others have a strong incentive not to keep their promises, they are unlikely to keep them either. The normative expectations attendant on the old norm can still be in place, making people fear that they would endure sanctions by keeping their word. Through deliberation, an exogenous agent can challenge these normative beliefs, paving the way for an easier transition to a new norm. A new norm can also be favored by the development of normative expectations, potentially accompanied by sanctions for norm violators. Here again, discussions and deliberations should make an important contribution. A punishment scheme that is devised through discussion is perceived as more legitimate, and a punishment accompanied by an explanation is more effective.

University of Pennsylvania

References

- Ackerman, B., & J. S. Fishkin. 2004. *Deliberation Day*. New Haven: Yale University Press.
- Aldashev, G., C. Imane, J.-P. Platteau, & Z. Wahhaj. 2010a. "Using the Law to Change the Custom." *Journal of Development Economics* 97(2): 182–200.
- Aldashev, G., J.-P. Platteau, & Z. Wahhaj. 2010b. "Legal Reform in the Presence of a Living Custom: An Economic Approach." *Proceedings of the National Academy of Sciences of the United States* 108(4): 21320–21325.
- Allport, F. 1924. *Social Psychology*. Boston: Houghton Mifflin Company.
- Balliet, D. 2010. "Communication and Cooperation in Social Dilemmas: A Meta-analytic Review." *Journal of Conflict Resolution* 54(1): 39.
- Benaquisto, L., & P. J. Freed. 1996. "The Myth of Inmate Lawlessness: The Perceived Contradiction between Self and Other in Inmates' Support for Criminal Justice Sanctioning Norms." *Law and Society Review* 30:481–511.
- Bicchieri, C. 2002. "Covenants without Swords: Group Identity, Norms, and Communication in Social Dilemmas." *Rationality and Society* 14(2): 192–228.
- Bicchieri, C. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Bicchieri, C., & A. Chavez. 2010. "Behaving as Expected: Public Information and Fairness Norms." *Journal of Behavioral Decision Making* 23(2): 161–178.
- Bicchieri, C., & Y. Fukui. 1999. "The Great Illusion: Ignorance, Informational Cascades and the Persistence of Unpopular Norms." *Business Ethics Quarterly* 9:127–155.

- Bicchieri, C. & A. Lev-On. 2007. "Computer-mediated Communication and Cooperation in Social Dilemmas." *Politics, Philosophy, and Economics* 62:139–168.
- Bicchieri, C., A. Lev-On, & A. Chavez. 2010. "The Medium or the Message? Communication Richness and Relevance in Trust Games." *Synthese* 176(1): 125–147.
- Bicchieri, C., & R. Muldoon. 2010. "Social Norms." *The Stanford Encyclopedia of Philosophy*. Retrieved from <http://plato.stanford.edu/entries/social-norms/>
- Brown, C. L. 2006. *Moral Capital*. Chapel Hill: University of North Carolina Press.
- Diop, Nafissatou, Modou Mbacke Faye, Amadou Moreau, Jacqueline Cabral, Helene Benga, Fatou Cisse, Babacar Mane, Inge Baumgarten, & Molly Melching. 2004. "The Tostan Program: Evaluation of a Community Based Education Program in Senegal." Population Council, Tostan.
- Elster, J. 1998. *Deliberative Democracy*. Cambridge: Cambridge University Press.
- Farrell, J., & M. Rabin. 1996. "Cheap Talk." *Journal of Economic Perspectives* 10:110–118.
- Festinger, L. 1957. *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.
- Fishkin, J. S., & R. C. Luskin. 2005. "Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion." *Acta Politica* 40(3): 284–298.
- Gillespie, D., & M. Melching. 2010. "The Transformative Power of Democracy and Human Rights in Nonformal Education: The Case of Tostan." *Adult Education Quarterly* 60:477–499.
- Gutmann, A., & D. F. Thompson. 1996. *Democracy and Disagreement*. Cambridge, Mass.: Belknap Press.
- Hadi, Amal Abdel. 2006. "A Community of Women Empowered: The Story of Deir El Barsha." In *Female Circumcision*, ed. Rogaia Mustafa Abusharaf. Philadelphia: University of Pennsylvania Press.
- Haidt, J., F. Bjorklund, & S. Murphy. 2000. "Moral Dumbfounding: When Intuition Finds No Reason." Unpublished manuscript. University of Virginia.
- Haile, Gabriel D. 2006. A Study on Community Dialogue Leading towards Abandoning of Harmful Traditional Practices, with Special Reference to Female Genital Cutting, Early Marriage, and Marriage by Abduction. UNICEF/Ethiopia.
- Hansen, K. M. 2003. *Deliberative Democracy and Opinion Formation*. University of Southern Denmark, Political Science and Public Management.
- Henrich, J., S. J. Heine, & A. Norenzayan. 2010. "The Weirdest People in the World." *Behavioral and Brain Science* 33(2/3): 1–75.
- Herrmann, B., C. Thöni, & S. Gächter. 2008. "Antisocial Punishment across Societies." *Science* 319(5868): 1362.
- Janssen, M. A., R. Holahan, A. Lee, & E. Ostrom. 2010. "Lab Experiments for the Study of Social-Ecological Systems." *Science* 328(5978): 613.
- Judd, C. M., & J. A. Krosnick. 1982. "Attitude Centrality, Organization, and Measurement." *Journal of Personality and Social Psychology* 42(3): 436.

- Kahan, D. 2000. "Gentle Nudges vs. Hard Shoves: Solving the Sticky Norms Problem." *University of Chicago Law Review* 67:607.
- Kamalski, J., L. Lentz, T. Sanders, & R. A. Zwaan. 2008. "The Forewarning Effect of Coherence Markers in Persuasive Discourse: Evidence from Persuasion and Processing." *Discourse Processes* 45(6): 545–579.
- Karpowitz, C. F., C. Raphael, & A. S. Hammond. 2009. "Deliberative Democracy and Inequality: Two Cheers for Enclave Deliberation among the Disempowered." *Politics and Society* 37(4): 576–615.
- Klassen, A. D., C. J. Williams, & E. E. Levitt. 1989. *Sex and Morality in the U.S.* Middletown, CT: Wesleyan University Press.
- Klofas, J. & H. Toch. 1982. "The Guard Subculture Myth." *Journal of Research in Crime and Delinquency* 19:238–254.
- Kurzban, R. 2001. "The Social Psychophysics of Cooperation: Nonverbal Communication in a Public Goods Game." *Journal of Nonverbal Behavior* 25(4): 241–259.
- LeJeune, L. & G. Mackie. 2009. *Social Dynamics and the Abandonment of Harmful Practices: A New Look at the Theory.* Unicef Innocenti Research Centre.
- Luskin, R. C., J. S. Fishkin, & R. Jowell. 2002. "Considered Opinions: Deliberative Polling in Britain." *British Journal of Political Science* 32(3): 455–487.
- Mackie, G. 1996. "Ending Footbinding and Infibulation: A Convention Account." *American Sociological Review* 61(6): 999–1017.
- Mansbridge, J. 1999. "Everyday Talk in the Deliberative System." In *Deliberative Politics*, ed. S. Macedo, 211–240. New York: Oxford University Press.
- Matza, D. 1964. *Delinquency and Drift.* New York: John Wiley and Sons.
- Mercier, H. 2011. "What Good Is Moral Reasoning?" *Mind & Society* 10(2): 131–148.
- Mercier, H., & H. Landmore. 2012. "Reasoning is for Arguing: Understanding the Successes and Failures of Deliberation." *Political Psychology* 33(2): 243–258.
- Mercier, H., & D. Sperber. 2011. "Why Do Humans Reason? Arguments for an Argumentative Theory." *Behavioral and Brain Sciences* 34(2): 57–74.
- Miller, D. T. & C. McFarland. 1991. "When Social Comparison Goes awry: The Case of Pluralistic Ignorance." In J. Suls and T. A. Wills (eds.), *Social Comparison: Contemporary Theory and Research.* Hillsdale, NJ: Erlbaum.
- O’Gorman, H. J. 1975. "Pluralistic Ignorance and White Estimates of White Support for Racial Segregation." *Public Opinion Quarterly* 39:313–330.
- Orbell, J. M., A. J. C. Van de Kragt, & R. M. Dawes. 1988. "Explaining Discussion-induced Cooperation." *Journal of Personality and Social Psychology* 54(5): 811–819.
- Ostrom, E. 1991. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge: Cambridge University Press.
- Ostrom, E., J. Walker, & R. Gardner. 1992. "Covenants with and without a Sword: Self-governance is Possible." *The American Political Science Review*, 404–417.

- Platteau, J.-P. 2000. *Institutions, Social Norms and Economic Development*. London: Harwood Academic Publisher.
- Packard, J. S., & D. J. Willower. 1972. "Pluralistic Ignorance and Pupil Control Ideology." *Journal of Education Administration* 10:78–87.
- Robinson, C. E. 1932. *Straw Votes*. New York: Columbia University Press.
- Sally, D. 1995. "Conversation and Cooperation in Social Dilemmas." *Rationality and Society* 7(1): 58.
- Schank, R. L. 1932. "A Study of Community and Its Group Institutions Conceived of as Behavior of Individuals." *Psychological Monographs* 43(2): 1–133.
- Stuntz, W. 2000. "Self-Defeating Crimes." *Virginia Law Review* 86:1871–1882.
- Sunstein, C. R. 2002. "The Law of Group Polarization." *Journal of Political Philosophy* 10(2): 175–195.
- Tetlock, P. E., R. S. Peterson, C. McGuire, S. Chang, & P. Feld. 1992. "Assessing Political Group Dynamics: A Test of the Groupthink Model." *Journal of Personality and Social Psychology* 63(3): 403–425.
- Thagard, P. 2002. *Coherence in Thought and Action*. Cambridge: The MIT Press.
- Ullmann-Margalit, E. 1977. *The Emergence of Norms*. Oxford: Oxford University Press.