The Psychology of Social Norms and the Promotion of Human Rights

Deborah A. Prentice

Princeton University


This chapter was written while I was Visiting Faculty in the School of Social Sciences at the Institute for Advanced Study, Princeton, NJ. I would like to thank Jeremy Adelman, JoAnne Gowa, Bob Keohane, Eric Maskin, Dale Miller, Catherine Ross, Teemu Ruskola, Rick Shweder, and Eric Weitz for helpful discussions and comments on earlier drafts of the chapter.

Please direct correspondence to:
Deborah Prentice
Department of Psychology
Princeton University
Green Hall
Princeton, NJ 08540
predebb@princeton.edu
Promoting human rights means changing behavior: Changing the behavior of governments that mistreat suspected criminals, opponents of their policies, supporters of their political rivals, and members of particular gender, ethnic, or religious groups; changing the behavior of corporations that mistreat their workers, damage the environment, and produce unsafe products; and changing the behavior of citizens who mistreat their spouses, children, and neighbors. In this chapter, I consider what an understanding of how social norms function psychologically has to contribute to this very worthy project. Social norms have proven to be an effective mechanism for changing health-related and environmental behaviors, so there is good reason to think that they might be helpful in the human-rights domain as well.

In the social sciences, social norms are defined as socially shared and enforced attitudes specifying what to do and what not to do in a given situation (see Elster, 1990; Sunstein, 1997). They are one source of constraint on behavior, in a family with laws, markets, and structural features of the environment (Lessig, 1998). Social norms are similar to laws, only less formal: The standards they promote are not codified, and their enforcement occurs informally, through social sanctions within groups and communities. They are similar to markets, only the incentives involved are social, not material: the approval of one’s friends and neighbors is what is at stake. Social norms are akin to structural features of the environment, in that they are social constraints that channel behavior in some directions and prevent it from going in other directions. One might assume that because social norms are less explicit and observable than laws, markets, and physical structures, they are also less powerful; in fact, the opposite is the case. Social norms do not require legislation, and the resources used to enforce them are unlimited; thus, they are much more pervasive than laws and markets. Moreover, because humans are highly social creatures, they are keenly sensitive to the behavior of those around them and are often prepared to put their
material and even physical well-being aside in pursuit of social approval (Miller & Prentice, 1994). The fact that social sanctions are freely available does not diminish their power.

Traditionally, advocates of human rights have conceived of their project strictly within a legal framework, narrowly defined. However, as the scope of that project has broadened to include the civil, economic, and social rights of particular groups of people within societies, the importance of social norms has drawn their attention. For scholars, norms provide a way of conceptualizing both sides of the human rights equation: the practices that violate people’s rights and the international standards that seek to uphold those rights (see, e.g., Checkel, 1999; Clark, 2001; Cortell & Davis, 1996; Risse, Ropp, & Sikkink, 1999; Simmons, 2000). For practitioners, norms are a highly attractive target for intervention, especially for rights violations that are embedded in the everyday practices of ordinary people. Of course, in all of these cases, the promise of social norms stems, in large part, from the fact that more formal and resource-dependent mechanisms are not available or suitable. Norms seem to offer a simple, economical way to explain, create, and sometimes unravel behavioral regularities.\(^1\) Exactly how they work remains a bit of a mystery. Let me begin the task of demystifying norms by outlining how they function to regulate behavior. Here, I will adopt a psychological level of analysis to examine how norms affect the behavior of individuals.

*The Mechanics of Social Norms*

As I noted, social scientists typically define norms as socially shared and enforced attitudes specifying what to do and what not to do in a given situation. Psychologists, by

---

\(^1\) The apparent simplicity of norms rests, at least in part, on their circularity. Sociologists, in particular, define social norms in terms of behavioral regularities and then posit norms as the cause of those regularities (Turner & Edgley, 2005). This is simple, to be sure, but not very productive. Defining norms as psychological entities – as mental representations – is one way to get around this problem.
contrast, favor a definition that is broader, more abstract, and more intrapsychic: Social norms are representations of where one’s group is located or ought to be located on an attitudinal or behavioral dimension (Miller & Prentice, 1996). These representations tend to be socially shared, and the attitudes and behaviors they specify tend to be socially enforced, but neither of these properties is definitional; neither is required of a norm. Instead, these are outcomes of the way norms function. Moreover, norms do not just specify what people ought and ought not to do; they also specify what people actually do, what they think, and how they feel. This definition of norms highlights a number of features that prove useful for understanding how norms work.

First, social norms are representations of a group, one’s own group; they are characterizations of what “we” think feel, and do. Thus, for there to be a norm, there must be a group, a “we,” to characterize. This point is important and often overlooked. The power of social norms to constrain behavior resides in social psychological dynamics that occur within groups – in the tendency of group members to look to each other for guidance, affirmation, and approval, and in the psychological pressures toward uniformity within the group that result (Turner, 1991). The group can be a club, an organization, a community, a neighborhood, a network; it can be loosely or tightly formed, formal or informal, face-to-face or virtual. What is important is that the members of the group consider themselves to be peers – equals, fellows, similar in important ways – and therefore consider the opinions and behaviors of the other group members to be relevant to their own.

This feature of social norms – their embeddedness in the dynamics of peer groups – begins to delimit the rights-relevant behaviors that norms can illuminate. For example, social norms might shape the interrogation practices used by a police force: Police officers use the behavior of their peers as a guide to what constitutes an appropriate level of force. Similarly,
social norms might influence the level of education girls receive in a community: Parents use the behavior of other parents as a guide. Social norms can explain uniformities of behavior among employees within corporations and among corporations within industries. They can also explain why groups differ – why some police forces use more aggressive interrogation tactics than others, some communities treat girls better than others, and some corporations and industries treat workers better than others.

For other behaviors, social norms probably play a less significant role. Consider, for example, the rights-related policies of governments; are these policies shaped by international human rights norms? This question has generated considerable controversy within the field of international relations, and has produced a growing literature on how, when, and where norms matter (see Cardenas, 2004 for a review). What the present analysis can contribute to that project is not answers, but more questions. One key question, for example, is whether the government has a peer group – that is, a group of similar others, whose behavior is relevant for comparison and whose good opinion is valued. If so, then the comparison and reinforcement processes within the international peer group might have a significant influence on adherence to human rights standards independent of other incentives. Another key question is whether the nation has internalized the international norms – that is, whether the nation itself now holds these standards as important for its own functioning. In this case, human rights norms might be enforced not by the international community alone, but also by domestic constituencies, and therefore may loom much larger in the government’s political calculus. The point I want to underscore is that the influence of norms begins and ends with membership in groups, and identifying the relevant groups is a critical step in understanding the nature of that influence.
Finally, on this issue, it is important to emphasize that social norms are particular to one’s own group (the ingroup). They are distinct from representations of other groups; those are stereotypes. They are also distinct from the thoughts, feelings, and behaviors of the individual members of the group. An individual may represent her group as holding attitudes and engaging in practices that she herself does not condone. For example, we may be in favor of the death penalty, and yet I may be vehemently opposed.

A second point to make about social norms is that they come in two types: Norms can be injunctive, characterizing where a group should be located on a dimension, or descriptive, characterizing where the group is located on the dimension (Miller & Prentice, 1996; Cialdini, Kallgren, & Reno, 1991). This is the difference between behavior in principle and behavior in practice, a difference that looms very large in the human-rights domain. In principle, a police force may prohibit aggressive interrogation tactics; in practice, officers may use these tactics when all else fails. In principle, a firm may have a norm of zero-tolerance for sexual harassment; in practice, suggestive remarks and sexist humor may be commonplace. Descriptive norms can be consistent with, inconsistent with, or independent of injunctive norms. Both descriptive and injunctive norms produce behavioral regularities, the former because violations of them are seen as odd and the latter because violations of them are seen as bad.

A third point to make about norms is that they are representations – ideas about the group, pictures people have in their heads about what their group does and values. These pictures have two distinguishing features. One, the group is represented as having a central tendency – an average or modal location on the dimension in question. For example, in the U.S., most people graduate from high school (82% of those who complete the 8th grade); most of those who do, enroll in a 4-year college (58%); and most of those who enroll go on to graduate (59% by age 26;
see Goldin & Katz, 2008). Thus, it is the norm in the U.S. to take an educational path that results in a bachelor’s degree; it is typical, expected. Of course, not all members of the group conform to this norm – indeed, the majority do not. But enough do that the central tendency is a meaningful feature of how the group is represented by its members. The central tendency is the content of the norm and determines the direction of its influence.

A second feature of the representation of the group is its dispersal – that is, how uniform the group is seen as being on the dimension in question. For example, suppose in one community, children receive, on average, 8 years of education, with a range from 7 to 9 years. In another community, children also receive an average of 8 years of education, but here the range is 4 to 12 years. The central tendency of the group, and therefore the direction of the norm’s influence, is the same in these two communities; however, the strength of that influence is likely to be very different. In the first community, children’s education is highly uniform; most children are treated the same way. In the second community, the range is so great that almost anything goes. As a consequence, the norm for how much to educate children is likely to have a much stronger influence in the first community than in the second.

Finally, it is important to emphasize that representations do not always map perfectly onto reality. People sometimes have ideas about their group that are outdated, exaggerated, or just plain wrong. In particular, under certain circumstances, people misjudge the location of the social norm, overestimate the uniformity of the group, and overestimate the extent of private support for group norms. These errors are of potential interest to human rights advocates for at least two reason: One, they often play a role in perpetuating extreme, anomalous, and dysfunctional group behaviors that interfere with people’s rights. In some cases, they foster these behaviors; in others, they inhibit people from openly expressing disapproval of these behaviors.
Two, erroneous assumptions about the group give rise to discrepancies between injunctive and descriptive norms, on the one hand, and private attitudes, on the other – discrepancies that, when brought to light, can catalyze behavior change. I will have more to say about these misrepresentations and their consequences in the sections that follow.

Social norms enter into discussions of human rights in two main ways: One, as a source of problematic behavior, and two, as a way to counter that behavior. Let me analyze how norms function in each of these capacities.

**Social Norms as a Source of Problematic Behavior**

Many of the behaviors that human rights advocates find problematic are enshrined in cultural practices and supported by (locally) injunctive norms. Female genital cutting is one widely discussed example, but there are many others: male circumcision, corporal punishment of children, infanticide, animal sacrifices, slavery, polygamy, and so on. Debates about these kinds of behaviors typically revolve around the tension between promoting human rights (conceived as universal, inalienable, and natural), on the one hand, and respecting local cultural practices, on the other. However, as Shweder (2006) has pointed out, local cultural practices are protected by rights as well – among them, the right of groups and nations to autonomy, self-determination, and freedom of religion, and the right of parents to raise their children within their own cultural traditions. Therefore, so long as the behavior in question remains a cultural norm – that is, so long as it remains prevalent, valued, and linked to important social outcomes within the group – it is very difficult for human rights advocates, coming from outside the group, to find legitimate grounds to overturn it.

There is, however, reason to question the actual degree of support any particular cultural norm enjoys. As I noted in the foregoing section, norms are representations that do not always
map perfectly onto reality. They are subject to systematic biases, and one of those biases is a tendency for people, both within and outside of the group, to overestimate the uniformity of support for normative practices. This bias in representation is rooted in three well-documented biases in the information people reveal to their fellow group members the inferences they draw from that information.

First, people are more likely to reveal their normative behaviors – those that reflect what is considered typical and desirable for group members – and to conceal their counternormative, and therefore stigmatizing, behaviors. College students talk about their drinking escapades and remain silent about Saturday nights spent at the student center; juvenile delinquents brag about acts of vandalism and conceal acts of obedience; residents in vegetarian co-ops eat tofu with their housemates and meat only with their close friends (Kitts, 2003; Miller & Prentice, 1996).

Second, people are more likely to express opinions they believe to be consistent with descriptive and injunctive group norms than those they believe to be inconsistent with these norms. Public opinion researchers have long recognized that believing oneself to be deviant is silencing, even when that belief is inaccurate (Noelle-Neumann, 1984). Empirical studies have confirmed that the willingness of group members to voice their opinions is a direct function of how prevalent they believe those opinions to be among their peers (Glynn, Hayes, & Shanahan, 1997). When the relationship between voice and perceived support for one’s position does not hold, it is either because people mistakenly believe that everybody agrees with them (Miller & Morrison, in press) or because they see their opinions as violating the descriptive but not the injunctive norm of the group – that is, they see themselves as different but good (Morrison & Miller, 2008). For example, in a group that is mildly in favor of capital punishment, a rabid capital punishment supporter may recognize that his opinion is deviant, but because he is on the
right side of the issue, he will feel different in a good way, superior to those who take a more moderate view. Thus, he will not hesitate to give voice to his opinion.

Third, people assume that their observations of their fellow group members – the behaviors they witness, the opinions they hear expressed – reflect what those individuals privately think, feel and do. That is, they fail to recognize how, and how much, people’s public behaviors and expressions are influenced by social constraints (Gilbert, 1998; Ross, 1977). The uncertainties and misgivings that people often feel about the norms of their groups, and their reluctance to challenge those norms under any and all circumstances, go unrecognized in all but the self. This tendency to take other people’s behavior at face value leads individuals to think that they feel differently than others do, even when they act identically. Researchers have termed this phenomenon pluralistic ignorance (see Prentice & Miller, 1996).

In short, systematic biases in the information people reveal to each other and the inferences they draw from that information can conspire to produce widespread overestimation of private support for social norms. Under these conditions, norm-challenging interventions run much less risk of infringing on people’s rights, not to mention desires. A key question, then, is when does overestimation of support for social norms occur? That is, when does public behavior diverge systematically from private views?

One condition that contributes to this divergence is constraint on public behavior. Of course, norms always constrain public actions, but sometimes the constraints are more explicit than others. When constraints are very strong and explicit, public actions and private views typically diverge quite substantially. For example, in countries living under totalitarian rule, such divergences are commonplace. However, this is not the situation that leads to an overestimation of private support for public practices. More insidious are those cases in which constraints on
public behavior are subtle and accompanied by the illusion of choice; these are the cases in which norms are often thought to have more support than they do. In his excellent book on divergences between public actions and private views, Kuran (1995) noted that he developed his theory of preference falsification not through experience with despotic regimes, but instead in societies with free press and open debate about official policies. Open societies can be deceptive in this regard: The constraints on behavior imposed by social roles, social norms, and public opinion go unrecognized, and people’s actions are seen by any and all observers as a straightforward reflection of their private preferences.²

An example of this phenomenon, one with considerable relevance to human-rights concerns, comes from studies of prison guards and inmates (Kauffman, 1981; Klofas & Toch, 1982; Toch & Klofas, 1982; Wheeler, 1961). Behavior by members of both of these groups is highly constrained by social roles. The traditional role of the guard demands harsh and punitive behavior toward inmates, and the traditional role of the prisoner demands resistant and recalcitrant behavior toward guards. It is very common for the behavior of both groups to polarize along this adversarial axis. Guards assume that other guards hold anti-inmate attitudes that correspond to their antagonistic stance, and prisoners assume that other prisoners hold anti-guard attitudes that correspond to their antagonistic stance. In fact, opinions surveys have shown that attitudes on both sides are much more congenial and sympathetic to the other group and, in fact, that both sides feel beleaguered by their peers’ hostility. Because they do not recognize the constraints introduced by their polarized roles in this situation, they also do not recognize that the feelings of their peers, like their own, are more moderate.

² For many reasons, norms-based approaches are likely to have little utility for combating rights violations committed by entrenched totalitarian regimes. However, they can be used to undermine the legitimacy of such regimes, as Kuran’s (1995) analysis suggests.
Constraints on public behavior are one source of a divergence between public actions and private views; a second is a lack of determinacy in people’s feelings and preferences. Such indeterminacy arises, for example, when preferences are not well-formed, when people do not know how to act on what they want, when they have mixed feelings, and when the situation itself is unclear or changing rapidly. Under all of these circumstances, people find themselves without a clear, internal guide to action; their inclination is to hesitate, to wait for clarity to emerge, and to look to other people for guidance. When everybody does this simultaneously, one of two things happens: Either inaction itself emerges as the norm, or inaction is interpreted as support for the status quo.

A famous example of the emergence of inaction as a social norm comes from research on bystander intervention. Bystanders in emergency situations are notorious for failing to act, failing to intervene. In the 1960s and 1970s, two researchers, Bibb Latané and John Darley, did a series of laboratory and field experiments that demonstrated that the likelihood that a witness to an emergency will do something to help is inversely proportional to the number of witnesses present (see Latané & Darley, 1970). In their analysis of the phenomenon, Latané and Darley highlighted the uncertainty inherent in emergency situations and the role that trying to cope with that uncertainty publicly plays in the dynamic. As they put it: “Occasionally the reactions of others provide false information as to the true state of their feelings… If each member of a group is trying to appear calm and not overreact to the possibility of danger, each other member, in looking to him for guidance, may be misled into thinking that he is not concerned. Looking at the apparent impassivity and lack of reaction of the others, each individual is led to believe that nothing is really wrong” (Latané & Darley, 1970, p. 44). In short, inaction, which is originally driven by uncertainty, becomes a norm as everybody observes it in each other. Note that people
in these situations need never decide anything; often, they are still uncertain, still waiting for clarity to emerge, when the situation ends. Moreover, time is not on their side: The longer they wait, the less likely they are to act (Latané & Darley, 1970).

An interesting version of this dynamic occurs during times of social change, when existing norms are being questioned and contested. It is common, under these circumstances, for attitudes to change more quickly than norms do, as people are persuaded by arguments for change but are not sure how or whether to act on their evolving convictions. As they look to others for clarity and guidance, they see only continued support for the status quo, support that they, in their inaction, are signaling as well. In other words, each individual’s inability or reluctance to act differently maintains the illusion of widespread support for norms to which people no longer feel allegiance. This phenomenon is known as a conservative lag (Fields & Schuman, 1976).

The Civil Rights Movement in the United States engendered a well-documented case of a conservative lag. National surveys conducted in the late 1960s revealed a sizable gap between the views people expressed about racial segregation and the beliefs they held about other people’s views. For example, the Survey Research Center’s 1968 National Election Study revealed that explicit support for racial segregation among whites was quite low across the country, ranging from just under 10% in the New England and Pacific states to just over 30% in the Southern states. By contrast, nearly 50% of white respondents expressed the belief that all or most whites in their area favored segregation (O’Gorman, 1975). Estimates of others’ support covaried with the observed degree of segregation within and across regions, suggesting that

---

3 Note that there is ample reason to take a skeptical stance toward self-reported racial attitudes, as they are subject to various forms of bias, some that people are aware of and some that they are not. However, in the case of the 1968 NES survey, additional evidence supported the validity of the self-reported attitudes, at least in regions outside the South (see O’Gorman 1975).
respondents based these estimates on what they saw around them. Fields and Schuman (1976) reported a similar result from a 1969 survey of Detroit residents. Asked whether they thought a white mother should allow her young daughter to bring a black classmate home from school, 76% said yes. Asked how “most people in Detroit” would answer, just 33% said yes. Again, predictions of others mirrored the high degree of racial segregation that characterized life in Detroit in the late 1960s. Why did people respond differently for themselves? Fields and Schuman (1976) argued strongly for the veracity of these self-reported attitudes. They maintained that the majority of whites in Detroit truly believed that a white mother should allow her young daughter to bring a black classmate home from school. These whites did not act on this belief because they, and everyone else, black and white alike, thought it was a minority view.

Surveys of racial attitudes no longer reveal a conservative lag; however, that is not to say that norms of racial segregation are a thing of the past. In many settings, some of them engineered to be racially diverse, norms of segregation continue to characterize behavior. One example of such a setting is the modern-day college campus, where racial and ethnic groups co-exist side by side, and students interact much more with ingroup members than with outgroup members. When asked why they do not interact more across racial boundaries, students typically say that they would like to have more contact, that they are not content with the status quo. However, they believe that members of other groups are not interested in having more contact with them. Black and white students alike attribute their own failure to cross group boundaries to a fear of being rejected, but they attribute the other group’s failure to cross group boundaries as reflecting a genuine lack of interest (Shelton & Richeson, 2005). It is easy to see how this dynamic would be self-perpetuating.
One final example illustrates how this misinterpretation of inaction can lead to an overestimation of citizen support for government policies. In the period between September 11th and the invasion of Iraq, there was considerable uncertainty over how the U.S. should conduct its War on Terror. U.S. citizens knew how they felt – they were sure they were terrified – but they did not know how to translate that fear into action. The U.S. government, as it turns out, did not know what to do either, and what emerged was a policy whereby the U.S. would decide what it wanted to do, would seek approval from the international community, and, if approval was not forthcoming, would take unilateral action. One of the results of that policy was the invasion of Iraq in 2003.

In the weeks leading up to the invasion, Todorov and Mandisodza (2004) conducted a survey using a representative national sample, in which they asked, “What do you think is the more important lesson of September 11: that the U.S. needs to work more closely with other countries to fight terrorism or that the U.S. needs to act on its own more to fight terrorism?” Respondents reported their own opinions and then estimated the proportion of Americans who endorsed each of the stated positions in the question. The results showed that private attitudes were overwhelmingly multilateral, with over three-quarters of respondents endorsing multilateralism as the more important lesson of September 11. However, these same respondents indicated that they thought unilateralism was the view of a majority of other Americans. The survey also asked specifically about attitudes toward a unilateral invasion of Iraq, without the support of the UN Security Council. Among respondents who expressed multilateral views, responses to this question depended on what they thought others thought. Specifically, when they thought others held unilateral views, they were more likely to agree with a unilateral invasion and more likely to support one even if they disagreed than when they thought others held
multilateral views (Todorov & Mandisodza, 2004). In short, the majority’s mistaken assumption that others supported government policies led them to support these policies themselves.

There is a different, but related, question one might ask about the gap between social norms and private views that emerged in the wake of September 11: Why did the government’s policies and the norms they spawned fail to represent the private views of citizens? Why were they so extreme? Most answers to these questions foreground particular features of the Bush Administration: Its power structure, ideology, history, decision-making processes, and so on. Alternatively, it seems worth exploring the possibility that there were also general forces at work, forces that tend to promote the emergence of extreme norms in highly uncertain or unstable situations. What might those forces be? Consider the qualities of individuals who are likely to act in an uncertain or unstable situation: The ones who will rush into a burning building, be the first in their neighborhood to invite a black classmate over to play, and think they know what to do in the face of a terrorist attack. These are, at the very least, highly unusual, unrepresentative people. They are certain when others are not; express their views when others remain silent; and take action while others are still wondering what to do. Sometimes their actions are heroic, and sometimes they simply make a bad situation worse. Importantly, they are a source of new, often extreme, and at least initially unsupported social norms. Norms promoting excessive alcohol use, extreme political views, and delinquent gang behavior have all been traced to the actions of a vocal and unrepresentative minority (see Korte, 1972; Matza, 1964; Miller & Morrison, in press).

In summary, social norms do not always reflect what the individuals who abide by and enforce them consider desirable and appropriate behavior; indeed, they often diverge markedly from individual feelings and preferences. Conceptualizing social norms as representations of the
group clarifies how and when these divergences occur. Their implications for understanding and promoting human rights are considerable: They help to explain why ordinary people stand by while atrocities occur, why groups in adversarial relationships tend to polarize, why customary practices remain entrenched even after individuals have been persuaded of their disutility, and why groups and governments often take extreme actions in times of uncertainty or crisis. Moreover, these situations also offer to human rights advocates opportunities to intervene without the usual concerns about infringing on the rights and sensibilities of those they seek to help. To the extent that there is, within the group, unacknowledged disaffection for existing norms and a majority sentiment for change, promoting that change becomes a much more manageable project.

**Social Norms as a Mechanism for Behavior Change**

I now turn to a consideration of what this analysis of social norms and how they function psychologically has to contribute to efforts at behavior change, and specifically, change in the behaviors that violate human rights. For human-rights advocates, the target of behavior change is typically the descriptive norm; that is, the goal is to change what the group is doing. One might want to move the location of the descriptive norm: to increase the average number of years girls go to school in a community, for example, or to reduce the incidence of circumcision or corporal punishment or spousal abuse. One also might want to change the dispersion of the group: to reduce the dispersion by tightening conformity to environmental or public health standards, for example, or to increase the dispersion by liberating people to express unpopular political views or to interact with whomever they choose. In the former case, one seeks to bring the descriptive norm into line with the injunctive norm; to get people to practice what they preach. In the latter
case, one seeks to weaken or eliminate the descriptive norm so that people will be guided by personal attitudes and predilections or perhaps by other norms or incentives.

There are two complementary schools of thought on the social psychology of changing behavior. One focuses on changing people’s attitudes and beliefs in order to change their actions (cf. Hovland, Janis, & Kelley 1953). For adherents to this school of thought, changing behavior is fundamentally about changing minds; it is about controlling the information to which people are exposed and how they process that information. This tradition survives and is most prominently represented in the fields of communications, marketing, and consumer behavior (see, e.g., Hornik, this volume). A second school of thought focuses not on attitudes and beliefs, but on motives – on the pressures pushing people in the direction of some goals and away from others (cf. Lewin, 1951). For adherents to this school of thought, changing behavior is fundamentally about changing the force of the different pressures, so as to change the equilibrium of the tension system (see Ross & Nisbett, 1991). This tradition survives and is most prominently represented in the fields of group dynamics, organizational behavioral, and behavioral economics.

Social norms fit comfortably within this latter tradition. Social norms are motivational pressures acting on all members of a group simultaneously to produce collective movement toward some behaviors and away from others. Behavior change, then, involves strengthening or weakening the force of particular social norms so as to change the equilibrium of the system. Given that the force of norms comes from the perception that everyone agrees with and adheres

---

4 Note that, as representations of the group, social norms do tap into attitudes and beliefs: In particular, beliefs about what the group thinks is the right and reasonable thing to do. As such, communication programs and information campaigns often have a key role to play in changing these representations and therefore in changing the direction and strength of their influence. Indeed, such campaigns can be central to the effort to create the consensus around which norms can form.
to them – that is, from the perceived dispersion of the group around the norm – the most straightforward way to strengthen or weaken them is to alter perceptions of that dispersion: to highlight the fact that everyone wants the same thing or that they all want different things. Note that this type of intervention often has the effect of changing the location of the norm as well. Specifically, strengthening a norm usually changes its location because behavior is not symmetrically distributed around the central tendency to start with (e.g., people are more likely to pollute more than the norm prescribes, not less, so strengthening the norm moves the location of the descriptive norm in the direction of less pollution); similarly, weakening a norm does not send behavior in all directions equally, but rather predominantly in the direction counter to the force of the norm (see Katz & Schanck, 1938 for a discussion of the dynamics of the J-curve).

The other point to make about this formulation is that it highlights the fact that social norms are part of a larger motivational system. They do not exist in isolation, but rather form systems of complementary and competing pressures on behavior. Moreover, that system includes not just social norms but other types of motives as well, including needs, goals, values, and material incentives; all of these forces act in concert to determine behavior. Thus, the consequences of strengthening or weakening a particular norm depend on what other forces are operating within the system.

*Strengthening Social Norms*

Strengthening social norms involves leveraging human sociality to mobilize conformity to a particular standard. Often, this is exactly what human rights advocates want to do: They want individuals, groups, corporations, and nations to adhere to the human-rights standards they have agreed are desirable; they want people to practice what they preach; they want to close the gap between descriptive norms and injunctive norms. Conformity turns out to be reasonably
straightforward to produce: All one needs to do is to highlight the direction and uniformity of the group’s behavior, and group members will follow suit. Of course, that only serves to strengthen existing behavioral patterns and only works in the short-run. Human rights advocates typically aspire to a more ambitious project: They seek to produce long-lasting change in the behavioral patterns of the group. This project runs into two challenges.

One challenge is to make the injunctive, rather than the descriptive, norm the salient guide for behavior. In many cases of human-rights violations, people agree on the desirable and appropriate course of action; they simply fall short. They respond to short-term incentives, give other goals and norms priority, and react badly in times of crisis. The group’s behavior ends up deviating systematically and predictably from the injunctive norm. How to close the gap? One solution is to enhance the salience and therefore the force of injunctive norms by highlighting the group’s collective approval or disapproval of a certain behavior. This potential effectiveness of this strategy was demonstrated by Schultz and colleagues (2007). The injunctive norm these researchers sought to promote was energy conservation. They partnered with a local utilities company in San Diego, which gave them access to the company’s routine meter readings of household energy consumption. The researchers gave householders feedback on their energy consumption at two time points, by way of a handwritten note delivered on a door hanger. For one set of households, they provided information about that household’s energy consumption and the average consumption of households of their size, with tips for how to conserve; this was the descriptive-norm intervention. For a second set of households, they provided the same information, plus a hand-drawn happy face if that household’s consumption was below average and a sad face if that household’s consumption was above average; this was the descriptive+injunctive-norm intervention. Both of these interventions induced conformity: When
presented with only descriptive information, above-average households decreased their subsequent consumption and below-average households increased their consumption. However, when presented with descriptive and injunctive information, above-average households decreased their subsequent consumptions and below-average households showed no change. Behavior tightened around the injunctive norm. This is exactly the outcome that human-rights advocates so often seek. Utilities companies were impressed with this outcome also: Several companies have now begun to include social comparison information, complete with smiling and frowning faces, in the design of their utility bills (Kaufman, 2009).

A second challenge is to maintain the salience of one particular norm amidst competing motivational pressures. Although people readily conform to social norms, both descriptive and injunctive, the motivational potency of those norms depends critically on their salience (Cialdini et al., 1991). And given the very crowded condition of most people’s motivational landscape, the salience of any particular norm waxes and wanes. These two facts have led some theorists to argue that what the field needs is a theory of norm salience, and in particular, its situational determinants (Bicchieri, 2000). Absent such a theory, and with a more pragmatic goal in mind, I would simply note that sustained behavior change, at a collective level, rarely rests on mobilizing conformity to one particular injunctive norm. That is, shaming rarely works in isolation. Rather, strengthening injunctive norms is an important part of a multifaceted strategy to produce behavior change, a strategy that typically combines the recruitment of directional social pressure with material and legal incentives. The success of the campaign to reduce smoking behavior over the last fifty years provides an excellent illustration of the importance of social norms in such a strategy (Brandt, 2007).
Weakening Social Norms

Of course, social norms are very often the problem, not the solution, to promoting human rights. That is, often human rights advocates seek to weaken the influence of norms that support rights-violating behaviors. The strategy here is the flip side of the strategies for strengthening norms. Specifically, if highlighting the direction and uniformity of the group’s behavior serves to strengthen norms, then decreasing the salience of the group’s behavior and/or highlighting its heterogeneity will serve to weaken them.

One way to implement this latter strategy is to give individuals voice: That is, to provide them with an opportunity to express their private views and to learn the private views of others. To the extent that prevailing norms do not have strong consensual support, this process will make salient to people the diversity of opinions within the group and thereby dispel the illusion of universality. Public opinion surveys can serve this function, sometimes inadvertently and occasionally despite the best efforts of those trying to conserve the norms. The history of prohibition laws in the U.S. provides a good illustration. As Katz and Schanck (1938) described: “During the prohibition era the forces in favor of prohibition never wanted any objective check on public sentiment. They tried to kill off straw polls on the subject of prohibition. As a consequence of their tactics even the politicians were fooled by an illusion of universality of opinion in favor of the Eighteenth Amendment. When an objective check was made, however, prohibition collapsed like a punctured balloon” (p. 175).

Liberating voice can also serve as an effective intervention strategy, particularly when social norms have less private support than people believe them to have. Schroeder and Prentice (1998) used just such a strategy to liberate college students from the norm of excessive alcohol consumption. Their intervention capitalized on an earlier finding that college students’
systematically and substantially overestimate their peers’ comfort with heavy drinking (Prentice & Miller, 1993). In the norm-focused intervention, students saw evidence that they themselves overestimated their peers’ comfort with heavy drinking and then discussed, as a group, the phenomenon of pluralistic ignorance and what might give rise to it in drinking situations. This intervention produced a 40% drop in drinking behavior, relative to a control intervention, at a follow-up assessment four to six months after the discussions. Additional findings suggested that this reduction in drinking was a result of weakening the power of the drinking norm. First, there was no evidence for a change in the location of the norm; that is, perceptions of the average student’s comfort with drinking did not vary as a function of the intervention, nor was there any evidence that a change in the perceived location of the norm mediated drinking behavior. Thus, students were not simply conforming to a lower drinking norm. Second, the intervention was most effective at reducing heavy drinking among students high in social anxiety, suggesting a reduction in perceived peer pressure. Third, the intervention left non-drinking students feeling more comfortable on campus, again suggesting a reduction in perceived peer pressure (see Prentice & Miller, 1996). All of these finding point to a weakening of the norm’s influence as a consequence of the intervention.

Liberating people from oppressive group norms may sound appealing, but it has at least two potential downsides. First, weakening norms deprives people of something they very much want and need: a way to fit in with their peers and fellow community members. Group norms are double-edged swords: They thwart individuality, but at the same time, provide identity and belongingness (Prentice, 2006). If people are deprived of one way to connect themselves to the group and establish themselves as good group members, they will find another way to do it. Or, they may seek out like-minded peers, who share their perceptions and habits, with whom they
can establish more well-grounded norms. Or, they may revert to the discredited norm once the
glow of liberation has worn off. The point is that intervention efforts may be able to liberate
people from a particular norm, but they will not be able to eliminate normative influence
altogether. Indeed, this would be neither a desirable nor a feasible goal. Second and relatedly, the
ultimate success of this type of intervention, from a human rights perspective, rests entirely on
the question of what will guide people’s behavior once the influence of the norm is eliminated.
As problematic as the norm may seem in prospect, there is always the possibility that what
comes afterward will be considerably worse.

Economic incentives are another effective way to weaken social norms, though they, too,
have a downside for human rights advocates. Specifically, economic taxes and subsidies work by
replacing moral considerations with material ones (Bowles, 2008); they have the effect of
removing behavior from the domain of right and wrong. Consider, for example, the case of blood
donation. In the early 1970s when Britain was considering trying to increase its blood supply by
moving from a strictly charitable system to a compensated system, Richard Titmuss (1971), a
social policy authority, argued that this move would have adverse effects on the both the quantity
and quality of the blood supply. Specifically, in a very controversial book, he argued that
compensating blood donation would make it less attractive to people because the compensation
would deprive them of the gratification they received from what heretofore had been an act of
civic virtue. In other words, he argued that the offer of an economic subsidy would shift the
behavior from a moral economy to a market economy. Moreover, he maintained that the
magnitude of the economic subsidy would not be high enough even to maintain current levels of
blood donation, much less to increase them. Titmuss (1971) offered little data in support of his
claims; however, recent theory and evidence have borne them out, at least to some extent
(Mellstrom & Johanneson, 2008). For blood donation, as well as for any other behaviors linked to injunctive norms, the psychological benefit people receive from the action is undermined if they begin to receive economic benefit from the action as well.

Economic taxes and fines work the same way. Consider the findings of an intervention study conducted in 10 private Day Care centers in Israel (Gneezy & Rustichini, 2000). The targeted group was the parents of children enrolled at each center, many of whom were routinely tardy to pick up their children. The intervention strategy was one of traditional deterrence: The investigators sought to reduce the frequency of tardiness by imposing a fine on late-coming behavior. At 6 of the day-care centers, a modest fine was imposed in the 5th week of the 20-week observation period and was removed in the 17th week; the other 4 centers served as a control group. The results showed that the number of late pick-ups increased significantly with the imposition of a fine and remained at the increased level even after the fine was removed. The investigators interpreted these results as evidence for the de-moralizing effects of economic incentives. Before the fine, behavior was regulated by injunctive norms against late-coming, to which parents imperfectly conformed. They tried to arrive on time and felt guilty when they did not make it. Once a fine was introduced, however, late-comers no longer felt guilty; they simply paid a material price for their lateness. The fine changed the meaning of the behavior – it made late-coming a commodity, rather than a transgression – and this new meaning persisted even once the fine was removed.

Note that the (ostensible) goal of this intervention was to mobilize conformity to the injunctive social norm; to bring parents’ behavior into line with the no-tardiness standard. The foregoing discussion suggests that a more effective way to accomplish this goal might be to strengthen the timeliness norm, perhaps by giving parents feedback about their own frequency of
late-arrivals compared with the late-arrivals of other parents, with frowning faces meted out to anyone who exceeds the average and happy faces to those who fall short of it. Posting the names of the most frequent latecomers might be another way to mobilize conformity to the injunctive norm; however, it would be important not to allow to list to grow too long, lest it communicate the commonness of the practice.

In short, economic incentives, both subsidies and fines, weaken social norms, but they do so by taking behavior off the moral grid. In the two examples I described, the introduction of economic incentives, designed to increase a desirable behavior, were insufficient even to maintain it. Of course, this is not always the case: When fines and subsidies are set high enough, they are usually sufficient to produce the desired behavior. Moreover, when the relevant actors are institutions, corporations, or governments – entities that are often more responsive to their material interests than to the good example and opinion of their peers – the use of economic incentives to promote human rights is essential. Nevertheless, it seems unlikely that the project of promoting human rights among ordinary people will be advanced by taking behavior out of the realm of right and wrong.

The question, then, for human rights advocates is: How can we free people up from the pressure of a particularly problematic social norm without weakening or even challenging the broader normative framework? That is, rather than liberating behavior from social control or relegating control to market forces, can interventions reduce or eliminate particular practices by changing in more subtle ways the balance of motivational forces within the normative system? The answer is yes: Legal remedies, wisely deployed, can be effective in this regard. Consider, for example, the case of dueling in the Southern United States (Lessig, 1995). Eradicating dueling in the South proved an elusive goal for many decades. Dueling was a normative means of resolving
disputes or matters of honor for Southern gentlemen. Because of the strong injunctive norms promoting dueling, participation in duels was often rational for the duelists; it enhanced their social status. However, dueling was costly to society at large, or at least it was perceived as costly by the states that struggled to ban the practice. Steep legal penalties for dueling, sometimes including death, had little deterrent value. One problem was the difficulty of maintaining enforcement of the law. Another was the strength of the norm – the high social cost incurred by refusing the challenge to a duel. No material tax on conviction for dueling could compete with the social (not to mention psychological) cost of being branded a coward and no gentleman. Moreover, Southern males of high social standing could not use the illegality of the behavior as a reason not to duel because the code of honor was meant to be above laws made by commoners.

A more effective solution, as it happened, was to make disqualification from holding public office one of the penalties for a dueling conviction (Lessig, 1995). This strategy served to liberate gentlemen from one powerful social norm by pitting an even more powerful norm against it: the obligation of the Southern gentlemen to be willing to hold public office. Conflicted duelers were unable to fulfill that obligation. Thus, refusing to duel could now be viewed as consistent with, or at least not entirely inconsistent with, the code of honor, the normative framework within which this class lived. In short, the law functioned not so much to constrain Southern males from dueling as to liberate them from having to duel in order to prove their allegiance to the norms of their group. Note that this strategy could only have worked if there were a desire among the majority of group members to extract themselves from social pressure to engage in this practice. Southern gentlemen must have wanted a way out, or this particular penalty would not have been effective. One can only assume that dueling either never had
widespread support or that support had waned over time and a conservative lag still held the norm in place (Fields & Schuman, 1976).

The history of the Southern United States provides many other examples of how the introduction of laws was effective because those laws liberated people from particular social norms to which they no longer felt allegiance without alienating them from society (Lessig, 1995). In some of the most striking cases, the members of the group themselves actually advocated for the laws that would force them to change their behavior. One such case is the role that Southern businesspeople played in the 1960’s Civil Rights legislation. According to Lessig (1995), in the hearings surrounding this legislation, many white business, restaurant, and hotel owners in the South testified in support of legislation that would make it illegal to do what they currently did – namely, refuse to serve African Americans. In other words, these individuals were advocating to be forced to do what they would not do voluntarily. Why didn’t they simply start serving African Americans on their own? Voluntarily serving African Americans would have been a deviant action; it would have opened them up to accusations that they were too greedy, or perhaps too sympathetic to blacks, and, in any case, that they put their own interests and desires ahead of the good of the community. It would have alienated them from their society. This social cost deterred them from doing something that was in their economic interest: to serve as many customers as they could. Anti-discrimination laws allowed them – indeed, forced them – to pursue their self-interest. In effect, they were lobbying for one kind of constraint to free them from a more costly one.

In short, laws can be an effective mechanism for liberating people from specific social norms without challenging the normative framework and without alienating people from their social group. Note that this strategy works only if people want to be liberated from the norms,
and even then, the sanctions imposed must be carefully titrated in light of prevailing (and also perceived) public opinion. New laws must fit into the broader normative framework; otherwise, they will not be enforced. Thus, laws prohibiting dueling in the South were ineffective because they challenged directly a broader normative framework that people were not prepared to relinquish: the code of honor. A more modest legal effort to target dueling in particular, while leaving this normative framework intact (indeed, capitalizing on it), was much more enforceable and therefore more effective (Lessig, 1995). More recent laws prohibiting date rape, drunk driving, and domestic violence have run up against the same problem. Although people agree that these undesirable behaviors should be reduced or eliminated, they are not prepared to accept legal remedies that challenges the prevailing normative framework. However, more modest efforts to stigmatize these behaviors from within the framework – often by construing them as violations of even more widely held and strongly supported injunctive norms – have proven quite effective. These successes may pave the way for more ambitious legal remedies in the future (see Kahan, 2000).

**Concluding Remarks**

Understanding the role that social norms play in regulating behavior can be a maddening enterprise: Norms are, at once, ubiquitous and elusive; they are both cause and effect, a source of their own perpetuation. Although the claim that norms play an important part in promoting and subverting human rights seems uncontroversial, it is a challenge to say anything more precise and useful than that. I have argued that a psychological level of analysis – a consideration of how people represent the norms of their groups and how those representations influence behavior at an individual level – can bring some measure of clarity to this enterprise. Moreover, the analysis
I have sketched suggests several conclusions about the most effective ways to design and deploy norms-based interventions in the service of promoting human rights.

One conclusion pertains to the conditions under which norms-based interventions are most likely to be effective for changing behavior: These interventions work only when there is some consensus around the desirability of change. That consensus can come from the injunctive norm (we are not doing what we should be doing) or from personal preferences (we are not doing what I want to be doing); it can be acknowledged or unacknowledged, explicit or latent. There simply needs to be some collective resonance to the notion that we would be better off if we did things differently. When there is, norms-based interventions can highlight the gap between the descriptive norm and either the injunctive norm or personal preferences to mobilize behavior change. When there is no gap – when people see the group as doing what it should do and what they personally want to do, at least on average – there is little scope for such an approach. In this latter case, intervention efforts need to start by changing people’s views of what is desirable, through education, information campaigns, or other kinds of consciousness-raising exercises.

A second conclusion concerns implementation. Social scientists have recently shown considerable interest in the question of how to produce behavior change and specifically how forceful an intervention to use. Current thinking suggests that less is more, that a nudge is often more effective than a shove (Kahan, 2000; see Thaler & Sunstein, 2008). This proposal resonates with an older psychological literature that contrasted the ineffectiveness of many highly ambitious and heavy-handed intervention efforts with the power of minor and often subtle manipulations of channel factors (see Ross & Nisbett, 1991 for a review). The analysis I have offered suggests a more complex and nuanced conclusion. On the one hand, it underscores the
fact that one should never underestimate the effectiveness of a well-placed nudge. For example, when it comes to liberating people from norms that no longer have private support, a modest intervention is often all that is required. In addition, a nudge is often all that is possible, particularly when the collective desire for change is weak or ambivalent (Kahan, 2000). On the other hand, sometimes behavior change requires a hard shove. Sometimes it requires a direct and forceful assault on behavior itself, to which injunctive norms and private attitudes then adjust. Laws mandating desegregation of public schools and those prohibiting smoking in public places are two examples of the hard-shove approach, where a forced change in descriptive norms was on the leading edge of changes in attitudes and injunctive norms. Of course, hard shoves, like gentle nudges, cannot be enacted without support, if not from the majority of the group then at least from its most powerful and influential members.

A final conclusion focuses on the broader picture. Social norms are just one part of an individual’s motivational system, a system that includes personal goals and values, social rules and norms, material incentives, legal constraints, and many other forces as well. This system can be perturbed intentionally by the types of psychological, material, and legal interventions I have outlined, but it also changes and evolves in response to shifting internal and external conditions over time. This gives the system an unpredictable quality; a strategy that works to change one behavior will not work to change another, a strategy that works in Group X will not work in Group Y, and a strategy that fails at Time A may very well succeed at Time B. Variations in outcome are almost always comprehensible in retrospect, but they are much more difficult to predict in prospect. This unpredictability should not discourage efforts at intervention – indeed, there is much that can be done – but it does highlight the fact that behavior change is an art, as well as a science. An understanding of the psychology of social norms, and of motivational
systems more generally, is vital to this project, but so too is particularistic knowledge of the behavior one seeks to change, the forces acting on that behavior in the status quo, the group that enacts the behavior, and the social, political, and legal climate surrounding the behavior. Only when they are armed with this local knowledge and scientific insight are behavior-change agents, including those who seek to promote human rights, in a position to assess whether, when, and how to intervene.
References


