

ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

UNLEASHED

Cass R. Sunstein

Discussion Paper No. 938

11/2017

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center

The Social Science Research Network Electronic Paper Collection:
<https://ssrn.com/abstract=3025749>

Unleashed

Cass R. Sunstein*

Abstract

Significant social change often comes from the unleashing of hidden preferences; it also comes from the construction of novel preferences. Under the pressure of social norms, people sometimes falsify their preferences. They do not feel free to say or do as they wish. Once norms are weakened or revised, through private efforts or law, it becomes possible to discover preexisting preferences. Because those preferences existed but were concealed, large-scale movements are both possible and exceedingly difficult to predict; they are often startling. But revisions of norms can also construct rather than uncover preferences. Once norms are altered, again through private efforts or law, people come to hold preferences that they did not hold before. Nothing has been unleashed. These points bear on the rise and fall (and rise again, and fall again) of discrimination on the basis of sex and race (and also religion and ethnicity). They also help illuminate the dynamics of social cascades and the effects of social norms on diverse practices and developments, including smoking, drinking, police brutality, protest activity, veganism, drug use, crime, white nationalism, “ethnification,” considerateness, and the public expression of religious beliefs.

I. “It’s Really Not A Problem”

In the late 1980s, when I was a visiting professor at Columbia Law School, I happened to pass, in the hallway near my office, a law student (female) speaking to an older law professor (male). To my amazement, the professor was stroking the student’s hair. I thought I saw, very briefly, a grimace on her face. It was a quick flash. When he left, I said to her, “That was completely inappropriate. He shouldn’t have done that.” Her response was dismissive: “It’s fine. He’s an old man. It’s really not a problem.”

Thirty minutes later, I heard a knock on my door. It was the student. She was in tears. She said, “He does this all the time. It’s horrible. My boyfriend thinks I should make a formal complaint, but I don’t want to do that. Please – I don’t want to make a fuss. Do not talk to him about it and do not tell anyone.”

* Robert Walmsley University Professor, Harvard Law School. I am grateful to Elizabeth Emens, Madeleine Joseph, Lucia Reisch, Daphna Renan, and Mark Tushnet for valuable comments on a previous draft.

Social norms imposed constraints on what the law student could say or do. She hated what the professor was doing; she felt harassed. But because of existing norms, she did not want to say or do anything. My little comment liberated her, at least in the sense that she felt free to tell me what she actually thought.

I am interested here in two different propositions. The first is that the erosion of social norms often unleashes people, in the sense that it *allows people to reveal what they believe and prefer, and to act as they wish*. New norms, and laws that entrench or fortify them, lead to the discovery of preexisting beliefs, preferences, and values. The discovery can be startling. The second is that *revisions of norms can construct preferences and values*. New norms, and laws that entrench or fortify them, can give rise to beliefs, preferences, and values that did not exist before.

Begin with the phenomenon of unleashing: When certain norms are in force, people falsify their preferences, or are silent about them, and as a result, strangers and even friends and family members may not be able to know about them.¹ Once norms are revised, people will reveal preexisting preferences and values, which norms had successfully suppressed. What was once unsayable is said, and what was once unthinkable is done.²

In the context of sexual harassment, something like this account is broadly correct: Women disliked being harassed, or even hated it, and revision of the norm was necessary to spur expression of their feelings and beliefs.³ (This account is incomplete, and I will complicate it.) Law often plays a significant role in fortifying existing norms, or in spurring their revision.⁴

¹ See Timur Kuran, *Private Truths, Public Lies* (1998).

² My focus is analytic, rather than normative, though a great deal might be said on the normative side. If, for example, one is roughly in accord with Mill's Harm Principle, see John Stuart Mill, *On Liberty* (1859), one might support norms that prevent harm to others; on a broadly compatible view, one might also be inclined to favor norms that solve coordination and collective action problems. See Edna Ullmann-Margalit, *The Emergence of Norms* (1978). Opposition to certain forms of discrimination would naturally lead to support for norms that discourage such discrimination. When and whether the force of law should be introduced in support of longstanding or emerging norms is another question. For relevant discussion, see Lawrence Lessig, *The Regulation of Social Meaning*, 62 *U Chi L Rev* 943 (1995). It should go without saying that some of the preferences, beliefs, and values that have been silenced by social norms deserve to be honored and given a spotlight, while others are destructive and hateful and so do not.

³ For a broadly consistent account, see Catharine MacKinnon, *Sexual Harassment of Working Women* (1978). Of course it is true that some harassers were and are women, and some objects of harassment were and are men. For simplicity of exposition, and to capture what mostly has happened and does happen, I bracket this point.

⁴ See Lawrence Lessig, *The Regulation of Social Meaning*, 62 *U Chi L Rev* 943 (1995).

Part of the importance of judicial rulings that forbid sexual harassment⁵ is that they revised norms. Whenever a new leader is elected, or whenever new legislation is enacted,⁶ it may have a crucial and even transformative signaling effect, offering people information about what other people think. If people hear the signal, norms may shift, because people are influenced by what they think other people think.⁷

But some revisions of norms, and laws that entrench those revisions, do not liberate anything. As norms begin to be altered, people come to hold, or to act as if they hold, preferences and values that they did not hold before.⁸ Revisions of norms, and resulting legal reforms, do not uncover suppressed desires; they produce new ones, or at least statements and actions that are consistent with new ones. With respect to men who sexually harassed women, that is a pretty good account, but it is incomplete, and I will complicate it.

Consider in this regard the idea of “political correctness,” which is standardly a reference to left-leaning social norms, preventing the expression of views that defy the left-of-center orthodoxy, and so silencing people.⁹ Political correctness means that people cannot say what they actually think; they are forced into some kind of closet. (The very term should be seen as an effort to combat existing norms.¹⁰) That is often what happens. But whether left or right, political correctness can also reconstruct preferences and values, making certain views unthinkable (for better or for worse). Eventually the unthinkable might become unthought.¹¹ Is that chilling? (Sometimes, but sometimes not; it is not terrible if no one thinks Nazi thoughts.)

Does it matter whether revisions of norms liberate people or instead construct new preferences and values? For purposes of understanding social phenomena, it certainly does. If preferences and values are hidden, rapid social change is possible and nearly impossible to

⁵ See *Meritor Savings Bank v. Vinson*, 477 U.S. 57 (1986); *Oncale v. Sundowner Offshore Services*, 523 US 75 (1998).

⁶ See Richard McAdams, *The Expressive Powers of Law: Theories and Limits* (2015), and in particular *id.* at 148: “When the perception of public attitudes falls seriously out of line with reality, legislators gain by enacting legislation corresponding to actual attitudes (and actual future votes), which produces a dramatic revelation – a ‘wake up call’ – of actual attitudes.”

⁷ See P. Wesley Schultz et al., *The Constructive, Destructive, and Reconstructive Power of Social Norms*, 18 *Psych. Science* 429 (2007).

⁸ Timur Kuran, *Ethnic Norms and Their Transformation through Reputational Cascades*, 27 *J. Legal Stud* 623 (1998).

⁹ A valuable discussion is Stephen Morris, *Political Correctness*, 109 *J Polit Econ* 231 (2001).

¹⁰ Part of the cleverness of the term is that it describes those who follow certain views as cowardly conformists, rather than people who are committed to hard-won principles. (That point is not meant to deny that the objections to “political correctness,” as a way of silencing dissenters, are often quite right. See Morris, *supra* note.)

¹¹ See Kuran, *supra* note.

predict.¹² When people are silent about their preferences or values, and when they falsify them, it is exceedingly difficult to know what they are. Because people conceal their preferences, outsiders cannot readily identify them.¹³ If people are discontent but fail to say so, and if they start to talk and act differently once norms are challenged and changed, then large-scale shifts in behavior are possible – but no one may have anticipated them.¹⁴ The rise of norms against sexual harassment is an example (which is hardly to say that sexual harassment has disappeared). The partial collapse of norms authorizing or promoting discrimination against transgender people can be seen in similar terms: For (many) transgender people, the effect is to prevent self-silencing and preference falsification. Similar dynamics help account for the fall of Communism,¹⁵ the Arab Spring,¹⁶ and the election of Donald Trump.¹⁷

When revisions in norms produce new preferences and beliefs, rapid change is also possible, but the mechanics are different. Those who produce such change do not seek to elicit preexisting preferences, beliefs, and values. As norms shift, people are not liberated. Influenced and informed by new or emerging norms, they develop fresh thoughts and feelings, or at least

¹² Timur Kuran, *Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution*, 61 *Public Choice* 41 (1989).

¹³ See Cristina Bicchieri and Yochitaka Fukui, *The Great Illusion: Ignorance, informational Cascades, and the Persistence of Unpopular Norms*, 9 *Business Ethics Quarterly* 127 (1999), and in particular this suggestion: “Norms whose existence depends on a collective illusion can be fragile. Whenever the veil of collective misperception is lifted, such norms may suddenly collapse.” *Id.*

¹⁴ See *id.*; Duncan Watts, *Everything Is Obvious* (2011); Susanne Lohmann, *The Dynamics of Social Cascades*, 47 *World Politics* 42 (1994). Note, however, that it would be possible to attempt to measure hidden preferences, beliefs, and values – for example, by guaranteeing anonymity. For an instructive finding that anonymity matters because of the role of social norms, see Leonardo Bursztyn et al., *From Extreme to Mainstream: How Social Norms Unravel* (2017), available at <http://www.nber.org/papers/w23415>, discussed below. For a demonstration that people are more willing to acknowledge discrimination online than on the phone, see *Expanding Discrimination Research: Beyond Ethnicity and to the Web*, Timur Kuran and Edward McCaffery, 85 *Social Science Quarterly* 713 (2004).

Even if we obtain a sense of hidden preferences, we might not be able to predict change, because of the importance of social dynamics, which cannot be anticipated in advance. See Susanne Lohmann, *I Know You Know He or She Knows We Know You Know They Know: Common Knowledge and the Unpredictability of Informational Cascades* 137, in *Political Complexity: Nonlinear Models of Politics* (Diana Richards ed. 2000); see also the discussion below.

¹⁵ Kuran, *supra* note.

¹⁶ Merouan Mekouar, *Protest and Mass Mobilization: Authoritarian Collapse and Political Change* (2016).

¹⁷ See Duncan Watts, *The Non-Inevitability of Donald Trump (and Almost Everything)* (2017), available at <https://medium.com/@duncanjwatts/the-non-inevitability-of-donald-trump-and-almost-everything-2a78e764183f>

act as if they have them.¹⁸ The rise of Nazism is famously complicated and highly disputed, but it can be understood in these terms.¹⁹

We can also find intermediate cases, in which people do not exactly have antecedent preferences that norms silence, but in which they hear a stubborn, uneasy voice in their heads, which they ignore, thinking, *Why bother to listen to that?* But as norms start to shift, that question has an answer: *Maybe it is telling me something important, or something that reflects my real feelings and beliefs.*

My principal examples involve discrimination, but the general points hold more broadly. Consider, for example, cigarette smoking, seatbelt-buckling, uses of green energy, purchases of organic food, considerateness,²⁰ demonstrations for white supremacy, veganism, the use of new languages,²¹ polyamory, religious practices,²² drug use, and crime. In all of these cases, norms can constrain antecedent preferences; new norms can liberate them or instead help construct new ones (or at least the appearance of new ones). In all of these cases, revisions in norms can result in large-scale changes in an astoundingly short time, including legal reforms, which can entrench and fortify those revisions.²³

II. What Discrimination Means

Begin with a working definition, offered by Jon Elster, who states that social norms are “shared by other people and partly sustained by their approval and disapproval. They are also sustained by the feelings of embarrassment, anxiety, guilt, and shame that a person suffers at the prospect of violating them.”²⁴ Because violations of social norms create such negative feelings, they impose costs on those who violate them. In that sense, they can operate in the same way as taxes,²⁵ which might turn out to be low or high. Importantly, however, some

¹⁸ Timur Kuran, *Ethnic Norms and Their Transformation through Reputational Cascades*, 27 *J. Legal Stud.* 623 (1998).

¹⁹ See Volker Ullrich, *Hitler: Ascent, 1889-1939* (2016). A highly controversial account is Daniel Jonah Goldhagen, *Hitler’s Willing Executioners* (1997); a counterpoint is Christopher Browning, *Ordinary Men* (rev. ed. 2017).

²⁰ See Edna Ullmann-Margalit, *Considerateness*, 60 *Iyyun* 205 (2011).

²¹ See Philip O’Leary, *The Prose Literature of the Gaelic Revival, 1881-1921* (2005).

²² See Anna Collar, *Religious Networks in the Roman Empire: The Spread of New Ideas* (2014).

²³ On law, Richard H. McAdams, *An Attitudinal Theory of Expressive Law*, 79 *Oregon Law Review* 339 (2000); McAdams, *supra* note. For an instructive account of the underlying social dynamics, see Susanne Lohmann, *I Know You Know He or She Knows We Know You Know They Know: Common Knowledge and the Unpredictability of Informational Cascades* 137, in *Political Complexity: Nonlinear Models of Politics* (Diana Richards ed. 2000).

²⁴ Jon Elster, *Rationality, Emotions, and Social Norms*, 98 *Synthese* 21, 23 (1994).

²⁵ I do not deal here with norms that solve collective action problems. For the defining account, see Edna Ullmann-Margalit, *The Emergence of Norms* (1977); see also Edna Ullmann-Margalit, *Normal Rationality* (2017). The analysis here could be adapted to that context, but with several

people are rebels, by nature or circumstance, and for them, defiance of social norms, taken as such, might be a benefit rather than a cost.²⁶

With respect to those who discriminate on the basis of sex and race (or other characteristics),²⁷ we can imagine four kinds of cases:

- (1) those in which discriminators want to discriminate, and norms allow them to do so;
- (2) those in which discriminators want to discriminate, but norms discourage or forbid them from doing so;
- (3) those in which discriminators do not want to discriminate, and norms allow them not to do so;
- (4) those in which discriminators do not want to discriminate, but norms encourage or require them to do so.

In cases (1) and (3), there is no conflict between preferences and norms. Case (2) is the familiar one; norms are operating as leashes or constraints. Note that in such cases, discrimination will not be observed, at least if the relevant norms are effective. Discriminators will falsify their preferences, or at least not reveal them. They will act as if they do not want to discriminate, even though they do. At the same time, they might hope to change the norm, at least in their community, and the question is whether they can succeed. To do so, they might well have to act collectively. Their efforts are far more likely to succeed if they are highly publicized.²⁸ (To be sure, some discriminators will simply defy the norm.²⁹)

wrinkles. The most important is that while such norms constrain people from doing as they wish (by preventing defection), they also operate to their mutual advantage, and so maintenance of the norm is in people's interest. For that reason, the leash that such norms impose should be quite welcome, at least on reflection.

²⁶ See Robert Kagan and Jerome Skolnick, *Banning Smoking: Compliance Without Enforcement*, in *Smoking Policy: Law, Politics, and Culture* (Robert L. Rabin ed. 1993), which shows that while most people comply with antismoking policies, even if they are not enforced, some do not, apparently because they like defying existing norms. The psychological finding of "reactance" is relevant here. See Sharon Brehm and Jack Brehm, *Psychological Reactance: A Theory of Freedom and Control* (1981).

²⁷ I focus on discrimination, because it is relatively tractable, but the analysis could be applied to any form of prejudice or hatred, even if it is manifested in something more extreme than discrimination (such as violence).

²⁸ Cristina Bicchieri, *Norms in the Wild* 164 (2017): "Even if trendsetters deviate from an established practice, they will never influence their peers if news of their deviance does not spread."

²⁹ See *id.* and in particular the treatment of the underlying character traits of "trendsetters."

There is an important role here for “norm entrepreneurs,”³⁰ operating in the private or public sector, who oppose existing norms and try to change them. Norm entrepreneurs draw attention to what they see as the stupidity, unnaturalness, intrusiveness, or ugliness of current norms. They may insist that many or most people secretly oppose them (and thus reduce pluralistic ignorance, understood as ignorance about what most people actually think³¹). Norm breakers – those who simply depart from existing norms, and refuse to speak or act in accordance with them – may or may not be norm entrepreneurs, depending on whether they seek to produce some kind of social change, or instead wish merely to do as they like.

In case (2), and all those like it, norm entrepreneurs might turn out to be effective, at least if the social dynamics, discussed below, work out in their favor.³² They might be able to signal not only their personal opposition to the norm, but also the existence of widespread (but hidden) opposition as well. The idea of a “silent majority” can be a helpfully precise way to signal such opposition.³³ Importantly, norm entrepreneurs might also change the social meaning of compliance with the norm: If they succeed, such compliance might suggest a lack of independence and look a bit pathetic, whereas those who defy the norm might seem courageous, authentic, and tough.

Case (4) may be the most interesting one. Here too, many discriminators will falsify their preferences. They will act as if they are sexist and racist, even though they are not. (In human history, that has often happened.) Faced with the stated conflict, what can discriminators do, short of defying the norm?

Here as well, norm entrepreneurs can act to alter the norm. They can also ask for or enlist law. Consider a revealing fact: *Some of the restaurants and hotels that that were regulated by the Civil Rights Act of 1964 actually lobbied vigorously for the legislation.*³⁴ Why –

³⁰ See Cass R. Sunstein, *Social Roles and Social Norms*, 96 *Colum. L. Rev.* 903 (1996); see Cristina Bicchieri, *supra* note, at 163-207 (2017). Both discriminators and objects of discrimination can of course serve as norm entrepreneurs. In the context of sexual harassment, Taylor Swift served as a norm entrepreneur in 2017, with her highly publicized objection to an act of unwanted touching. See Melena Ryzik, *Taylor Swift Spoke Up. Sexual Assault Survivors Were Listening* (Aug. 15, 2017), available at https://www.nytimes.com/2017/08/15/arts/music/taylor-swift-sexual-assault.html?_r=0

³¹ There are many accounts. For especially good ones, see Kuran, *supra* note; D. Garth Taylor, *Pluralistic Ignorance and the Spiral of Silence: A Formal Analysis*, 46 *Public Opinion Quarterly* 311 (1982). A valuable account, with special reference to law, is Richard McAdams, *The Expressive Powers of Law* 136-62 (2015).

³² See *infra*.

³³ See Elisabeth Noelle-Neumann, *The Spiral of Silence* (2d ed. 1993); D. Garth Taylor, *Pluralistic Ignorance and the Spiral of Silence: A Formal Analysis*, 46 *Public Opinion Quarterly* 311 (1982).

³⁴ See Lawrence Lessig, *The Regulation of Social Meaning*, 62 *U Chi L Rev* 943 (1995). See also Richard H. McAdams, *Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination*, 108 *Harv L Rev* 1003, 1065-85 (1995).

you might ask – would such companies affirmatively seek to be forbidden, by law, from discriminating on the basis of race? If they did not want to discriminate on the basis of race, they certainly could have stopped discriminating on the basis of race. Why did they need the law?

Norms help to explain what happened. The relevant companies had an antecedent preference: They wanted to make money. The best way to make money was to serve anyone who was willing to pay. For that reason, they did not want to discriminate. In fact they wanted *not* to discriminate, because discrimination, on their part, was costly. But in light of prevailing norms, they would incur a high cost for not discriminating, which would provoke a hostile reaction in their community. As Lawrence Lessig writes, “for a white to serve or hire blacks was for the white to mark him or herself as having either a special greed for money or a special affection for blacks.”³⁵ In these circumstances, the force of the law was needed to alter the social meaning of nondiscrimination. Once the Civil Rights Act of 1964 was enacted, nondiscrimination was a matter of compliance. Profit-making companies were liberated.

We can see related phenomena in other domains in which revised norms, or new law, work to counteract discrimination. No one should doubt that many men who have engaged in sex discrimination did not want to do so, in the sense that they acted in accordance with norms that they did not endorse (and might have abhorred). As norm entrepreneurs began their work, norms started to change, and as law prohibited discriminatory behavior, such men could do what they wanted to do. Of course this is far from a full picture of the consequence of new antidiscrimination norms. But it is part of it.

The phenomenon holds more broadly. Many people are glad that the law requires them to buckle their seatbelts, because it enables them to do as they wish. Many people support laws that forbid drug use, in part because such laws make it easier for them to decline to use drugs. Many people support drunk driving laws, in part because it enables them to decline to drive when they should not. When Ulysses took steps to resist the sirens, he did not enlist social norms, or seek laws that might alter the social meaning of actions, but some similarly situated people have been happy to be able to exactly that. Norms and norm revisions, and laws that codify them, can operate as precommitment strategies.

III. Preference Falsification and Adaptive Preferences

These are points about discriminators, who may not want to discriminate, and who can be liberated by new norms and by law that reflects or entrenches them. But as the Columbia law student story suggests, the most important effects of norms and norm revisions are likely to be on the objects of discrimination. In the simplest and most common cases, the objects of discrimination have an antecedent preference, and the norm prevents them from stating or

³⁵ Lessig, *supra* note, at 966.

acting on it.³⁶ The preference may even be falsified (as it was when the law student initially assured me that she did not object to what the professor was doing). In that respect, the objects of discrimination are like actors in a play; they are reciting the expected lines. In cases of sex and race discrimination, that is a familiar phenomenon.³⁷ The legitimation of the antecedent preference brings it out of the closet.

A. Norm Cascades and Multiple Equilibria

In circumstances of this kind, large-scale change is possible. Suppose that many people within a population object to discrimination, but because of existing norms, they do not say or do anything. Suppose that the objectors have different thresholds for raising an objection. A few people will do so if even one person challenges or defies the norm; a few more will do so if a few people challenge or defy the norm; still more will do so if more than a few people challenge or defy the norm; and so on. Under the right conditions, and with the right distribution of thresholds, a small spark can ignite a conflagration, eventually dismantling the norm.³⁸

It is important to emphasize that with small variations in starting points, and inertia, resistance, or participation at the crucial points, that may or may not happen. Suppose that a community has long had a norm in favor of discrimination on the basis of sexual orientation; that many people in the community abhor that norm; that many others dislike it; that many others do not care about it; that many others are mildly inclined to favor it; and that many others firmly believe in it. If norm entrepreneurs make a public demonstration of opposition to the norm, and if the demonstration reaches those with relatively low thresholds for opposing it, opposition will immediately grow. If the growing opposition reaches those with relatively higher thresholds, the norm might rapidly collapse. But if the early public opposition is barely visible, or if it reaches only those with relatively high thresholds, it will fizzle out, and the norm might not even budge. These are the two extreme cases. We could easily imagine intermediate cases, in which the norm suffers a slow, steady death, or in which the norm erodes but manages to survive.

It is for this reason that otherwise similar communities can have multiple equilibria, understood here as apparently or actual stable situations governed by radically different norms. After the fact, it is tempting to think that because of those different norms, the communities

³⁶ A classic discussion is William Festiner et al., *The Emergence and Transformation of Disputes: Naming, Blaming, Claiming . . .*, 15 *Law and Society Review* 631 (1980-81).

³⁷ Tocqueville, speaking of slavery, writes: "Is it a blessing of God, or a last malediction, this disposition of the soul that gives men a sort of depraved taste for the cause of their afflictions?" Quoted in Jon Elster, *Explaining Social Behavior* 199 (rev. ed. 2015). Tocqueville's comment is not of course an adequate account of slavery, but it vividly captures the phenomenon of adaptive preferences.

³⁸ The classic account is Mark Granovetter, *Threshold Models of Collective Behavior*, 83 *Am J Sociology* 1420 (1978); the idea is productively extended in Kuran, *supra* note.

are not otherwise similar at all, and to insist on some fundamental cultural difference between them. But that thought might be a product of an illusion, in the form of a failure to see that some small social influence, shock, or random event was responsible for the persistence of a norm in one community and its disintegration in another.³⁹

Some of the most interesting work on social influences involves the existence of informational and reputational “cascades”; this work has obvious relevance to the revision of norms and eventually legal reform.⁴⁰ A starting point is that when individuals lack a great deal of private information⁴¹ (and sometimes even when they have such information), they are attentive to the information provided by the statements or actions of others. If A is unaware whether genetic modification of food is a serious problem, he may be moved in the direction of alarm if B seems to think that alarm is justified. If A and B believe that alarm is justified, C may end up thinking so too, at least if she lacks independent information to the contrary. If A, B, and C believe that genetic modification of food is a serious problem, D will have to have a good deal of confidence to reject their shared conclusion. The result of this process can be to produce cascade effects, as large groups of people eventually end up believing something, simply because other people seem to believe it too. It should be clear that cascade effects may occur, or not, depending on seemingly small factors, such as the initial distribution of beliefs, the order in which people announce what they think, and people’s thresholds for abandoning their private beliefs in deference to the views announced by others.⁴²

Though the cascades phenomenon has been discussed largely in connection with factual judgments, the same processes are at work for norms; we can easily imagine norm cascades (information-induced or otherwise), which may well produce legal reform.⁴³ Some such

³⁹ Much science fiction is focused on this possibility. See Ray Bradbury, *A Sound of Thunder and Other Stories* (2005). For an empirical demonstration, see Matthew Salganik et al., *Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market*, 311 *Science* 854 (2006).

⁴⁰ See *id.*; Sushil Bikhshandani et al., *A Theory of Fads, Fashion, Custom, and Cultural Change As Informational Cascades*, 100 *J. Polit. Econ.* 992 (1992).

⁴¹ On the pervasiveness of this condition, see Russell Hardin, *How Do You Know? The Economics of Ordinary Knowledge* (2014).

⁴² For evidence, see Marc Willinger and Anthony Ziegelmeyer, *Are More Informed Agents Able To Shatter Information Cascades in the Lab*, in *The Economics of Networks: Interaction and Behaviours* 291 (Patrick Cohendet et al. eds.) (New York: Springer Verlag, 1996).

⁴³ An intriguing wrinkle is that when a cascade gets going, people might underrate the extent to which those who join it are reacting to the signals of others, and not their own private signals. For that reason, they might see the cascade as containing far more informational content than it actually does. See Erik Eyster and Matthew Rabin, *Naïve Herding in Rich-Information Settings*, 2 *American Economic Journal: Microeconomics* 221 (2010); Eric Eyster et al., *An Experiment on Social Mislearning* (2015), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2704746. Norm entrepreneurs have a strong interest in promoting this mistake.

cascades may be a product of information; some may involve values. Suppose, for example, that A believes that discrimination against transgender people is wrong, that B is otherwise in equipoise but shifts upon hearing what A believes, that C is unwilling to persist in his modest approval of discrimination against transgender people when A and B disagree; it would be a very confident D who would reject the moral judgments of three (apparently) firmly committed others. In such contexts, many people, lacking firm convictions of their own, may end up believing what (relevant) others seem to believe. Stylized as the example is, changes in social attitudes toward smoking, recycling, and sexual harassment have a great deal to do with these effects. And here as well, small differences in initial conditions, in thresholds for abandoning private beliefs because of reputational pressures, and in who hears what when, can lead to major differences in outcomes.

The availability heuristic⁴⁴ often plays a major role in norm cascades.⁴⁵ The basic idea is that judgments about probability are often made by asking whether relevant events come to mind. If, for example, a particular case of egregious discrimination receives a great deal of public attention, people might see, or come to believe, that such discrimination is widespread. In a variation on the availability heuristic, a single event might come to be highly salient, affecting not only probability judgments but also judgments about morality and norms. With respect to sexual harassment, Anita Hill's widely publicized allegations about Clarence Thomas had a significant effect on public perceptions.⁴⁶ Some people serve as "availability entrepreneurs"; they emphasize a single incident in an effort to create an availability cascade, involving facts or norms. In many contexts, the effects of civil disobedience (of norms or law) are greatly magnified by the unduly aggressive responses of official targets; those responses tend to be publicized, and they signal that those who engaged in disobedience may well have been right.⁴⁷

Thus far the discussion has emphasized purely informational pressures and informational cascades, where people care about what other people think because they do not know what to think, and they rely on the opinions of others, to show what it is right to think. But with respect to norms, the more important point is that there can be reputational pressures and reputational cascades as well. Because people care about their reputations, they speak out, or remain silent, or engage or decline to engage in expressive activity, partly in order to preserve those reputations, even at the price of failing to say what they really think. Suppose, for example, that A believes that climate change is an extremely serious environmental problem; suppose too that B is skeptical. B may keep quiet, or even agree with A, simply in order to preserve A's good opinion. C may see that A believes that climate change is a serious

⁴⁴ See Amos Tversky and Daniel Kahneman, Availability: A heuristic for judging frequency and probability, 5 *Cognitive Psychology* 207 (1973).

⁴⁵ See Timur Kuran and Cass R. Sunstein, Availability Cascades and Risk Regulation, 51 *Stan. L. Rev.* 683 (1999).

⁴⁶ See Amy Black and Jamie Allen, Tracing the Legacy of Anita Hill, 19 *Gender Issues* 33 (2001).

⁴⁷ See Edward Glaeser and Cass R. Sunstein, A Theory of Civil Disobedience (2015), available at <http://www.nber.org/papers/w21338>

problem, and that B seems to agree with A; C may therefore voice agreement even though privately she is skeptical or ambivalent.

It is easy to see how this kind of thing might happen in political life with, for example, politicians expressing their commitment to same-sex marriage or to eliminating discrimination against transgender persons (even if they are privately skeptical). Here too the consequence can be cascade effects – large social movements in one direction or another -- when a number of people appear to support a certain course of action simply because others (appear to) do so.

B. A “Down Look”

For discrimination, of course, it is too simple to say that its objects are opposed and silence themselves. When discrimination is widespread, and when norms support it, its objects might see it as part of life’s furniture. In some cases, they might not even feel that their preferences and values have been constrained. Consider Gordon Wood’s account of the pre-revolutionary American colonies, when “common people” were “made to recognize and feel their subordination to gentlemen,” so that those “in lowly stations . . . developed what was called a ‘down look,’” and “knew their place and willingly walked while gentlefolk rode; and as yet they seldom expressed any burning desire to change places with their betters.”⁴⁸ In Wood’s account, it is impossible to “comprehend the distinctiveness of that premodern world until we appreciate the extent to which many ordinary people *still accepted their own lowliness.*”⁴⁹

Wood urges that as republicanism took hold, social norms changed, and people stopped accepting their own lowliness.⁵⁰ His account is one of a norm cascade, but not as a result of the revelation of preexisting preferences. With amazement, John Adams wrote that “Idolatry to Monarchs, and servility to Aristocratical Pride, was never so totally eradicated from so many Minds in so short a Time.”⁵¹ David Ramsay, one of the nation’s first historians (himself captured by the British during the American Revolution), marveled that Americans were transformed “from subjects to citizens,” and that was an “immense” difference, because citizens “possess sovereignty. Subjects look up to a master, but citizens are so far equal, that none have hereditary rights superior to others.”⁵² Thomas Paine put it this way: “Our style and manner of thinking have undergone a revolution more extraordinary than the political revolution of a country. We see with other eyes; we hear with other ears; and think with other thoughts, than those we formerly used.”⁵³

⁴⁸ Gordon Wood, *The Radicalism of the American Revolution* 29 (rev. ed. 1998).

⁴⁹ *Id.* at 29-30.

⁵⁰ *Id.*

⁵¹ *Id.* at 169.

⁵² *Id.*

⁵³ Thomas Paine, “Letter to the Abbe Raynal,” in *Life and Writings of Thomas Paine*, Daniel Edwin Wheeler, ed. (New York: Vincent Parke and Company, 1908), 242. Quoted in Gordon Wood, *The Creation of the American Republic, 1776-1787*, revised ed. (Chapel Hill: University of North Carolina Press, 1998), 68.

Adams, Ramsay, and Paine appear to be speaking of new preferences, beliefs, and values, rather than the revelation of suppressed ones. How this happens remains imperfectly understood.⁵⁴ While the idea of preference falsification captures much of the territory I am exploring, it is complemented by situations in which adaptive preferences are altered by new or revised norms.

There are also intermediate cases, involving what might be called *partially adaptive preferences*. Objects of discrimination may not exactly accept discrimination. They might live with it, and do so with a degree of equanimity, thinking that nothing can be done. It is not a lot of fun to beat your head against the wall. In cases of partially adaptive preferences, objects of discrimination are not like actors in a play; they are not falsifying their preferences. Once norms change, some inchoate belief or value might be activated that was formerly suppressed, or that was like a small voice in the head. It is fair enough to speak of liberation, but the case is not as simple as that of the law student at Columbia.

IV. Liberating Isms

Some norms reduce discrimination, but others increase it. Suppose that people have antecedent hostility toward members of social groups; suppose that social norms constrain them from speaking or acting in ways that reflect that hostility. On one view, this is the good side of “political correctness”; it prevents people from expressing ugly impulses.⁵⁵ But norms that constrain sexism and racism are of course stronger in some times and places than in others, and they can be relaxed or eliminated. In the aftermath of the election of President Donald Trump, many people feared that something of this kind has happened (and are fearing that it continues to happen).⁵⁶ The basic idea is that President Trump is a norm entrepreneur; he is shifting norms in such a way as to weaken or eliminate their constraining effects. It is difficult to test that proposition in a rigorous way, but consider a highly suggestive experiment.

Leonardo Bursztyn of the University of Chicago, Georgy Egorov of Northwestern University and Stefano Fiorin of the University of California at Los Angeles attempted to test whether President Trump’s political success affects Americans’ willingness to support, in public, a xenophobic organization.⁵⁷ Two weeks before the election, Bursztyn and his colleagues recruited 458 people from eight states that the website Predictwise said that Trump was certain to win (Alabama, Arkansas, Idaho, Nebraska, Oklahoma, Mississippi, West Virginia and

⁵⁴ Kuran, *supra* note, is the best account.

⁵⁵ For relevant discussion, see Glenn Loury, *Self-Censorship in Public Discourse*, 6 *Rationality & Society* 428 (1994).

⁵⁶ For a journalistic account, see Jamelle Bouie, *What We Have Unleashed* (2017), available at http://www.slate.com/articles/news_and_politics/politics/2017/06/this_year_s_string_of_brutal_hate_crimes_is_intrinsically_connected_to_the.html

⁵⁷ See Leonardo Bursztyn et al., *From Extreme to Mainstream: How Social Norms Unravel* (2017), available at <http://www.nber.org/papers/w23415>

Wyoming). Half the participants were told that Trump would win. The other half received no information about Trump's projected victory.

All participants were then asked an assortment of questions, including whether they would authorize the researchers to donate \$1 to The Federation for American Immigration Reform, accurately described as an anti-immigrant organization whose founder has written, "I've come to the point of view that for European-American society and culture to persist requires a European-American majority, and a clear one at that."⁵⁸ If participants agreed to authorize the donation, they were told that they would be paid an additional \$1. Half the participants were assured that their decision to authorize a donation would be anonymous. The other half were given no such assurance. On the contrary, they were told that members of the research team might contact them, thus suggesting that their willingness to authorize the donation could become public.

For those who were not informed about Trump's expected victory in their state, giving to the anti-immigration group was far more attractive when anonymity was assured: 54 percent authorized the donation under cover of secrecy as opposed to 34 percent when the authorization might become public.⁵⁹ But for those who were informed that Trump would likely win, anonymity did not matter at all.⁶⁰ When so informed, about half the participants were willing to authorize the donation regardless of whether they received a promise of anonymity.⁶¹ The central point is that information about Trump's expected victory altered social norms, making many people far more willing to give publicly, and eliminating the comparatively greater popularity of anonymous endorsements.

As an additional test, Bursztyn and his colleagues repeated their experiment in the same states during the first week after Trump's election.⁶² They found that Trump's victory also eliminated the effects of anonymity -- again, about half the participants authorized the donation regardless of whether the authorization would be public.⁶³ The general conclusion is that if Trump had not come on the scene, many Americans would refuse to authorize a donation to an anti-immigrant organization unless they were promised anonymity. But with Trump as president, people feel liberated. Anonymity no longer matters, apparently because Trump's election has weakened the social norm against supporting anti-immigrant groups. It is now more acceptable to be known to agree "that for European-American society and culture to persist requires a European-American majority, and a clear one at that."

The central finding can be seen as the mirror image of the tale of the law student and the law professor. For a certain number of people, hostility to anti-immigrant groups is a

⁵⁸ Id.

⁵⁹ Id.

⁶⁰ Id.

⁶¹ Id.

⁶² Id.

⁶³ Id.

private matter; they do not want to voice that hostility in public. But if norms are seen to be weakening or to be shifting, they will be willing to give voice to their beliefs.

We can easily imagine much uglier versions of the central finding. When police brutality increases, when hateful comments or action are directed at members of certain religious groups, when white supremacy marches start, when ethnic violence breaks out, when mass atrocities occur, and when genocide is threatened, one reason is the weakening or transformation of social norms that once made the relevant actions unthinkable.⁶⁴ In some such cases, what was akin to a tax has been eliminated; in other cases, what was akin to a tax has been transformed into something like subsidy. The subsidy might be necessary to spur the destructive behavior, but for some participants, removal of the tax is enough.

IV. Internalized Norms

My emphasis has been on situations in which people have an antecedent preference or value, whose expression a norm blocks; revision of the norm liberates them, so that they can talk or act as they wish. (Sexual desires may be the most obvious example, where people may be startled to find out what they like, though in that context, there can be a complex interplay between discovery and construction of preferences.⁶⁵) But I have also noted that some norms are internalized, so that people do not feel chained at all. Once the norm is revised, they speak or act differently, either because they feel constrained by the new norm to do that, or because their preferences and values change. Orwell's *Nineteen Eighty-Four* is a chilling tale of something like that, with its terrifying closing lines: "But it was all right, everything was all right, the struggle was finished. He had won the victory over himself. He loved Big Brother."⁶⁶

That is the dark side. But return to the case of sexual harassment. Many men are appalled by the very thought of sexual harassment. They endorse, and do not feel constrained by, current norms. For them, norms and legal rules against sexual harassment are not a problem, any more than norms and legal rules against theft and assault are a problem. If they are older, some of these men might have experienced a shift over the course of their lives. If they are younger, some of these men might not be able to imagine a context in which sexually harassing someone would be a fun or good experience.

For such men, we do not have cases of preference falsification. For some of them, it might be clarifying to speak of adaptive preferences. But it is better to say that the relevant people are committed to the norm, so that defying it would not merely be costly; it would be unthinkable.

⁶⁴ Helpful discussion can be found in Timur Kuran, *Ethnic Norms and Their Transformation through Reputational Cascades*, 27 *J. Legal Stud* 623 (1998).

⁶⁵ The 2017 Netflix series, *Gypsy*, is a notable exploration of this topic.

⁶⁶ George Orwell, *Nineteen Eighty-Four* (1949).

Something similar can be said for many actions that conform to social norms. Most people are not bothered by nonexistence of a social norm against dueling.⁶⁷ For many people, seatbelt buckling and recycling are not properly characterized as costs; they are a matter of routine, and for those who buckle their seatbelts or recycle, the relevant actions may well be taken as a net benefit. When the social norm is one of considerateness,⁶⁸ those who are considerate usually do not feel themselves to be shackled; they want to be considerate. When this is so, the situation will be stable; norm entrepreneurs cannot point to widespread, but hidden, dissatisfaction with the norm. But for both insiders and outsiders, it will often be difficult to distinguish between situations in which norms are internalized and situations in which they merely seem to be. That is one reason that stunning surprises are inevitable.⁶⁹

⁶⁷ For an intriguing account of what happened, see Harwell Wells, *The End of the Affair? Anti-Dueling Laws and Social Norms in Antebellum America*, 5 *Vand. L. Rev.* 1805 (2001).

⁶⁸ See Ullmann-Margalit, *supra* note.

⁶⁹ On the possibility of learning about actual preferences, see note *supra*. Recall that another reason for unpredictability involves interdependencies among agents, which can produce changes that cannot be anticipated in advance. Susanne Lohmann, *I Know You Know He or She Knows We Know You Know They Know: Common Knowledge and the Unpredictability of Informational Cascades* 137, in *Political Complexity: Nonlinear Models of Politics* (Diana Richards ed. 2000). For evidence in a different but relevantly similar context, see Matthew Salganik et al., *Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market*, 311 *Science* 854 (2006).