



University of Kentucky  
UKnowledge

---

Law Faculty Scholarly Articles

Law Faculty Publications

---

Winter 2010

## A Behavioral Approach to Human Rights

Andrew K. Woods

*University of Kentucky College of Law*, [andrewkwoods@uky.edu](mailto:andrewkwoods@uky.edu)

Follow this and additional works at: [https://uknowledge.uky.edu/law\\_facpub](https://uknowledge.uky.edu/law_facpub)



Part of the [Human Rights Law Commons](#), and the [International Law Commons](#)

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

---

### Recommended Citation

Andrew K. Woods, *A Behavioral Approach to Human Rights*, 51 Harv. Int'l L.J. 51 (2010).

This Article is brought to you for free and open access by the Law Faculty Publications at UKnowledge. It has been accepted for inclusion in Law Faculty Scholarly Articles by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@lsv.uky.edu](mailto:UKnowledge@lsv.uky.edu).

---

## A Behavioral Approach to Human Rights

### Notes/Citation Information

Harvard International Law Journal, Vol. 51, No. 1 (Winter 2010), 51-112

## A Behavioral Approach to Human Rights

---

Andrew K. Woods\*

*For the last sixty years, scholars and practitioners of international human rights have paid insufficient attention to the ground level social contexts in which human rights norms are imbued with or deprived of social meaning. During the same time period, social science insights have shown that social conditions can have a significant impact on human behavior. This Article is the first to investigate the far-ranging implications of behavioralism—especially behavioral insights about social influence—for the international human rights regime. It explores design implications for three broad components of the regime: the content, adjudication, and implementation of human rights. In addition, the Article addresses some of the advantages and limitations of the behavioral approach and outlines the rich but unexplored nexus of behavioralism, norms, and international law.*

### INTRODUCTION

In 1946, the Nazi propagandist Julius Streicher was convicted and hanged at the International Military Tribunal in Nuremberg for persecution as a crime against humanity. The court held that his hateful words were a “poison Streicher injected into the minds of thousands of Germans which caused them to follow the National Socialist policy of Jewish persecution and extermination.”<sup>1</sup> The genocide, the court suggested, involved too large a swath of German society to be taken as just a series of bad acts committed by bad Germans; rather, it was a social phenomenon.<sup>2</sup>

Two years after Streicher’s hanging, the human rights movement was given international legal status with the Universal Declaration of Human Rights (“UDHR”). Drafted in direct response to the horrors of Nazi society, the UDHR rests on a very different assumption about human nature from the Streicher trial. Article 1 begins, “All human beings are born free and equal in dignity and rights. *They are endowed with reason and conscience . . .*”<sup>3</sup>

---

\* Gates Scholar, Cambridge University and Hauser Visiting Doctoral Researcher, New York University Law School. Special thanks to Ryan Goodman, Jack Goldsmith, and David Runciman for extraordinary feedback and guidance. Thanks also to Robert Goodin, Martha Minow, and Ganesh Sitaraman for helpful feedback; and to Albertina Antognini, Seth Ard, Jean-Denis Greze, and Matthew Perault for that and more. The students in the Spring 2008 International Law Workshop at Harvard Law School provided innumerable insights for applying social science work to international law; thanks to them as well. Any errors are mine alone.

1. 1 TRIAL OF THE MAJOR WAR CRIMINALS BEFORE THE INTERNATIONAL MILITARY TRIBUNAL: NUREMBERG, 14 NOVEMBER 1945–1 OCTOBER 1946, at 294–95 (1947) [hereinafter IMT].

2. See generally HANNAH ARENDT, EICHMANN IN JERUSALEM: A REPORT ON THE BANALITY OF EVIL 131–37 (1963) (noting that killing was so woven into German bureaucracy, it had become a banal metric for job performance).

3. Universal Declaration of Human Rights, G.A. Res. 217A, art. 1, U.N. GAOR, 3d Sess., 1st plen. mtg., U.N. Doc. A/810 (Dec. 12, 1948) (emphasis added).

The UDHR goes on to list the fundamental human rights in a series of articles, the majority of which are directed at individuals and accordingly begin with either “Everyone has the right to . . . .” or “No one shall be . . . .”<sup>4</sup>

The above texts—the Streicher trial’s declaration that German society had been infected and the UDHR’s emphasis on the individual’s capacity for reason—represent different assumptions about human behavior. The Streicher judgment implies that certain social conditions can drive human action, even unreasonable, unthinkable acts. Germans were “caused” to commit genocide by a “poison” that infected them; they did not individually will it.<sup>5</sup> The judgment is less concerned with the actual commission of individual acts of violence, or the punishment of those acts, than it is with the creation of a social situation that engendered great harm. Meanwhile, the text of the UDHR seems less concerned with social dynamics: “[a]ll human beings” are born with reason, and they “should act toward one another” so as to secure the individual rights outlined in the articles that follow.<sup>6</sup> The Streicher judgment recognizes the power of social conditions to influence behavior while the UDHR is grounded in the assumption that individual free will drives human action. In the language of today’s social psychology, the Streicher judgment contains a situationist understanding of human behavior; in the language of today’s economics it is behaviorist.<sup>7</sup> The UDHR, on the other hand, is concerned with rational unitary actors. The two texts imagine human rights violations emerging from different sources: social conditions and individual dispositions.

This is striking because while human rights scholarship and practice have focused overwhelmingly on unitary actors’ compliance with rules, social scientists have produced an abundance of insights about the power of social conditions to influence behavior.<sup>8</sup> Yet very little of that knowledge has been incorporated into human rights scholarship or practice.<sup>9</sup>

4. *Id.*

5. IMT, *supra* note 1.

6. Universal Declaration of Human Rights, *supra* note 3.

7. The behavioral notion of situationism and the research behind it is discussed in more detail in Part II. For now, both terms—behavioralism and situationism—will serve as placeholders for the many empirical insights about the power of social situations to influence individual behavior. This Article does not offer a definition of the terms for two reasons. First, the social science literature itself uses many different terms to describe the effects of our social surroundings on our behavior. More importantly, it would be unproductive to get bogged down in a definitional debate when the purpose of this Article is to show that regardless of the definitions one uses, human rights scholars and practitioners have tended to neglect behavioral insights and that these insights provide novel avenues for giving human rights norms meaning. Needless to say, the term “behavioral” is meant only to emphasize the claim that human rights law should be more concerned with behavior. The term is not meant to restrict this inquiry to “behavioral” economics at the expense of the historically richer fields of psychology or even the humanities.

8. See *infra* Part II.

9. There is considerable debate about which model of human action is more accurate. See generally Gregory Mitchell, *Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law*, 43 WM. & MARY L. REV. 1907 (2002). This Article does not enter that debate here—it

This Article is an early attempt to investigate the implications of recent behavioral insights—particularly the consequences of taking into account the regulatory power of social situations—for the international human rights regime as it exists today.<sup>10</sup> Scholars have used social theory to question the use of a rights framework to advance human rights norms.<sup>11</sup> This Article takes a different tack: it explores how human rights regime designers can utilize the latest research about social influence to enhance respect for human rights. Compared to the present regime, a behavioral human rights regime would rely less on classical assumptions about individual cognition and more on the known mechanisms of social influence. The result would be a greater focus on the potential for certain social conditions to generate human rights abuses and, inversely, for different conditions to promote human rights norms. Compared to present regimes, a behavioral human rights regime would be less concerned with rights violators and more concerned with environments conducive to rights violations. Under this approach, a human rights violation would be treated as a social phenomenon with distinct and identifiable social mechanisms, rather than a collection of bad actors committing a series of bad acts.

Human rights scholars and practitioners intuitively understand that context matters, but their models rarely take sufficient account of situations of social influence. Consider the following examples that illustrate the regulatory power of social context and how that context can give human rights norms meaning.

*Civilian Deaths in Afghanistan.*<sup>12</sup> In 2007, the non-governmental organization Human Rights Watch drafted a report addressing the rise in civilian casualties in the conflict in Afghanistan between North Atlantic Treaty Organization (“NATO”) and Taliban forces.<sup>13</sup> Skeptical that the Taliban would read the report, two researchers sent to the British Broadcasting Corporation (“BBC”) Pashto World Service audio interviews with Afghans who had lost family members to Taliban suicide bombings. The day after the broadcasts, the Taliban called the radio station, outraged, and denounced the report.<sup>14</sup>

---

is enough for present purposes to note that situations have an effect on behavior and that this effect is not currently explicitly accounted for by international legal scholars.

10. This Article uses the term “regulatory” in the broader socio-legal context of constraining action—not the narrower legal context of secondary legislation by the state.

11. See generally Duncan Kennedy, *The Critique of Rights in Critical Legal Studies*, in LEFT LEGALISM/LEFT CRITIQUE 178 (Wendy Brown & Janet E. Halley eds., 2002).

12. See *infra* Part II.B.

13. Human Rights Watch reports are usually fact-findings, heavily researched and often drafted and/or vetted by lawyers for the legal implications of the factual findings. Press releases follow. See generally Human Rights Watch, Reports, <http://www.hrw.org/en/publications/reports> (last visited Nov. 5, 2009).

14. See Telephone Interview with former HRW researchers Michael Sheikh and John Sifton (Sept. 22, 2009) (on file with Harvard Law Library).

*Hand Washing in Ghana.*<sup>15</sup> Communicable diseases are widespread in Ghana, despite the wide availability of soap. For years, health experts have tried to change hand-washing behavior through data-driven informational campaigns, to no avail. Last year, a British public health official, working with Unilever, launched a campaign to change the way people feel—rather than how they think—about bathroom use. In the targeted communities, the campaign appears to have worked: bathroom soap use is up.

*Facebook in Egypt.*<sup>16</sup> The government of Egypt does not grant its citizens freedom of speech or assembly, making protests both scarce and dangerous. During the last year, however, the size and number of protests directly critical of the government have increased substantially. One explanation is the rise of the social networking website Facebook.com, which is now the second most trafficked site in the country<sup>17</sup> and which features hundreds of groups critical of government policies. The explosion of protest groups, the theory goes, is the result of the fact that the social networking software revealed previously-hidden preferences about the Mubarak regime, creating the social conditions necessary to lend solidarity to would-be protestors.

*Detainee Treatment in Iraq.*<sup>18</sup> In January 2008, Colonel James Brown, the American Commander of Camp Bucca in southern Iraq, could be seen reaching down and picking up spent chewing gum from the gravel lot where Coalition Forces were holding over 20,000 detainees. Clutching the wad of chewed gum, he explained: “If I let one of my soldiers’ hair get out of place, I know abuse is not far behind.”<sup>19</sup>

These examples are telling. Norms, and the laws that codify them, are situationally contingent. They do not regulate behavior, as scholars have tended to assume, only through individual internalization: social conditions can activate norms that regulate behavior without any purposeful engagement from the individual whose behavior changes.<sup>20</sup> In each of the above cases, we see the prominence of the social situation as a regulator of behavior. In other, different situations, the relevant norms would not have had the desired effect on behavior; the difference in each case was social context. And

---

15. See *infra* Part I.C.

16. See Samantha M. Shapiro, *Revolution, Facebook-Style*, N.Y. TIMES MAG., Jan. 25, 2009, § 6 (Magazine) at 34 (noting that “for young people in Egypt, Facebook, which allows users to speak freely to one another and encourages them to form groups, is irresistible as a platform not only for social interaction but also for dissent”).

17. Alexa.com, Alexa Top 500 Global Sites, <http://www.alexa.com/topsites>. (last visited Nov. 11, 2009).

18. See *infra* Part III.A.

19. Interview with Col. James Brown, in Camp Bucca, Iraq (Jan. 9, 2008) (on file with the Harvard Law Library). See also Andrew K. Woods, *The Business End*, FIN. TIMES MAG., June 27, 2008, at 44.

20. See CRISTINA BICCHIERI, *THE GRAMMAR OF SOCIETY: THE NATURE AND DYNAMICS OF SOCIAL NORMS* 3 (2005) (“The simplistic, common view that we conform to norms either because of external sanctions or because they have been internalized flies in the face of much evidence that people sometimes obey norms even in the absence of any obvious incentive structure or personal commitment to what the norm stands for.” (citation omitted)).

yet, there has been no systematic attempt by legal scholars to think of social situations as a component of international legal regimes.

The idea that social situations are important components of a regulatory regime is not new to the fields of sociology, anthropology, or psychology, and it is an intuitive concept for many lawyers and judges.<sup>21</sup> Indeed, it has recently driven innovations in criminal jurisprudence.<sup>22</sup> But the idea has largely been under-appreciated by international legal scholars as a mechanism for engendering respect for human rights.<sup>23</sup> Just as the social norms literature showed over a decade ago, informal social mechanisms are an important part of the legal toolkit, especially in the context of human rights.<sup>24</sup>

Robert Cover famously wrote that law is not just a system of rules, but rather a “normative universe” made up of communities and their narratives that locate law and give it meaning.<sup>25</sup> That idea is powerful but incomplete. How does it function on the ground? What constitutes a “normative universe?” This Article builds on Cover’s idea with behavioral insights about social situations—potential normative universes—and applies it to international human rights regimes. Rather than ask, as most international human rights scholars have, whether or how human rights treaties influence state

21. See MAX HORKHEIMER & THEODOR W. ADORNO, *DIALECTIC OF ENLIGHTENMENT* 183–84 (1944) (reviewing the potency of Fascist propaganda and concluding that “[c]ommunication makes people conform by isolating them”).

22. See Adriaan Lanni, *The Future of Community Justice*, 40 HARV. C.R.-C.L. L. REV. 359, 364–80 (2005) (summarizing scholarship about and developments in the community justice movement, which uses social structures and norms to regulate community behavior in place of incarceration). See also Jeffrey Fagan & Tracey Meares, *Punishment, Deterrence and Social Control: The Paradox of Punishment in Minority Communities*, 6 OHIO ST. J. CRIM. L. 173, 180–85 (2008) (noting that a state’s system of legal control can crowd out its ability to create informal social systems of control); Dan Kahan & Tracey Meares, *The Coming Crisis of Criminal Procedure*, 86 GEO. L.J. 1153, 1167–70 (1998) (noting that anti-racism measures now limit the ability of racial minority communities to craft community policing measures).

23. In international relations, constructivists look to social situations to engender human rights treaty ratification, but they do not look to social situations to engender individual respect for human rights. Likewise, human rights scholars such as Sally Engle Merry discuss the contexts in which norm internalization takes place, but they do not treat social situations as formal units of the human rights regime. See *infra* Part II.

24. See, e.g., Lawrence Lessig, *The Regulation of Social Meaning*, 62 U. CHI. L. REV. 943 (1995). See also ERIC A. POSNER, *LAW AND SOCIAL NORMS* (2000). Describing his work applying social norms scholarship to the law, Lessig says: “My assumption is not that the tools that I describe do not exist or that they do not exist within economics. My claim is only that they need to become the ordinary tools of legal analysis.” Lawrence Lessig, *The New Chicago School*, 27 J. LEGAL STUD. 661, 677 (1998). The claim of this Article is similar: it is not that these are new social phenomena, but rather that situational insights should become the ordinary tools of human rights regime design.

25. “The student of law may come to identify the normative world with the professional paraphernalia of social control. The rules and principles of justice, the formal institutions of the law, and the conventions of a social order are, indeed, important to that world; they are, however, but a small part of the normative universe that ought to claim our attention. No set of legal institutions or prescriptions exists apart from the narratives that locate it and give it meaning. For every constitution there is an epic, for each decalogue a scripture. Once understood in the context of the narratives that give it meaning, law becomes not merely a system of rules to be observed, but a world in which we live.”

Robert M. Cover, *Foreword: Nomos and Narrative*, 97 HARV. L. REV. 4, 4–5 (1983) (footnotes omitted).

actors,<sup>26</sup> this Article looks at the local social contexts in which human rights norms are imbued with or deprived of social meaning.

The Article proceeds in four parts. Part One employs certain behavioral insights to explain the regulatory power of social situations, paying particular attention to four concepts: social proof, social networks, social meaning, and the social construction of identity.<sup>27</sup> Part Two then shows that human rights scholars and regime designers are insufficiently focused on social situations as components of their regulatory apparatus, and as a result underutilize recent knowledge about social influence. This insufficient focus on situation is evidenced by the scholarship's preoccupation with mechanisms for enforcing treaty commitments on recalcitrant states—an important project, but one that has paid too little attention to ways that international human rights law can influence human behavior—and by regime designers' preoccupation with exposing and sanctioning unitary rights violators whether they be states or individuals.

Part Three explores the implications of the above insights for human rights regime designers in terms of the content, the adjudication, and the implementation of human rights. Specifically, it investigates the implications of behavioralism for the way human rights law constructs the social conditions of its own compliance; debates about the scope of human rights crimes; the hierarchy of norms in human rights regimes; the mechanisms for delivering transitional justice; the development of human rights doctrine; and the role of new and existing avenues of social influence, including the role that media, education, and nonstate actors play in giving human rights social meaning.

The final Part outlines the benefits of this behavioral approach, and also acknowledges some of the difficult questions raised by the purposeful manipulation of the social medium, questions that encompass liberalism concerns and functional concerns. It also addresses concerns about the relevance of social situations as a factor in shaping individual behavior, as well as some of the potential challenges of usefully integrating attention to social situation into a formal regime.

## I. BEHAVIORAL INSIGHTS ABOUT SOCIAL INFLUENCE

Psychologists have repeatedly shown that people are heavily influenced by their social situations, even more than they realize. Behaviors that were once considered dispositional character traits have been shown to vary considerably under different situational conditions, raising doubts about the extent to which they are indeed character traits and not, in aggregate, traits that re-

---

26. See, e.g., Oona A. Hathaway, *Do Human Rights Treaties Make a Difference?*, 111 YALE L.J. 1935, 1938–39 (2002) (using a large-scale quantitative analysis to address the questions “Are human rights treaties complied with? Are they effective in changing states’ behavior for the better?”).

27. That is, social meaning to the people they are meant to protect. See *infra* Part I.C.



sult from a given situation.<sup>28</sup> In a now-famous series of experiments, Edward E. Jones and Victor A. Harris showed that humans tend to over-emphasize the role of dispositional factors, such as choice and preference, and under-appreciate the importance of situational factors,<sup>29</sup> when in fact, “[t]here is much evidence . . . that the most important determinant of one’s behavior is the situation in which one is.”<sup>30</sup> That is, not only are social situations a potent substrate for regulation, but, even more significantly, we are often blind to their regulatory effects.<sup>31</sup> The fact that humans are often unaware of their situations is one reason social regimes have such surprising regulatory power. Before applying this work to the realm of international human rights, it will be necessary to review some of the ways in which social situations function as regulatory regimes.

These social science insights do not share a unified nomenclature: psychologists speak of humans as “biased” in ways previously unknown;<sup>32</sup> economists say human action is sometimes “irrational”;<sup>33</sup> game theorists say that people do not play subgame perfect Nash equilibria;<sup>34</sup> political scientists and neuroscientists say that voters act on emotion more than reason, though they vocalize their preferences using the latter;<sup>35</sup> and so on. Still, by taking a broad scan of this behavioral work, a strong picture of the human actor emerges. This is an actor who—irrational or not—is sometimes self-regarding, sometimes other-regarding, and often a strong reciprocator; who belongs to multiple networks, each of which exert influence over him; who is subject to social norms, some of which are also network norms; who has a

---

28. See Lee D. Ross, *The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process*, 10 *ADVANCES IN EXPERIMENTAL SOC. PSYCHOL.* 173 (1977). See also Dan T. Gilbert & Patrick S. Malone, *The Correspondence Bias*, 117 *PSYCHOL. BULL.* 21, 21–38 (1995) (giving a history of attribution errors, and noting that Ross’s great insight was “that the social psychologist’s tendency to underestimate the power of situations (which he called the fundamental attribution error) was shared by social psychology’s subjects and that this was the key to understanding their behavior in a wide range of seemingly unrelated experiments.” *Id.* at 34).

29. See Edward E. Jones & Victor A. Harris, *The Attribution of Attitudes*, 3 *J. EXPERIMENTAL SOC. PSYCHOL.* 1, 4–14 (1967) (describing experiments in which subjects read pro-Castro statements and then rated the authors’ feelings toward Castro as sympathetic, even though subjects knew the authors had been asked to write such sympathetic statements as part of a psychological experiment). See also Edward E. Jones & Richard E. Nisbett, *The Actor and the Observer: Divergent Perceptions of the Causes of Behavior*, in *ATTRIBUTION: PERCEIVING THE CAUSES OF BEHAVIOR* 79 (Edward E. Jones et al. eds., 1972).

30. BICCHIERI, *supra* note 20, at 80.

31. See generally Jon Hanson & David Yosifon, *The Situation: An Introduction to the Situational Character, Critical Realism, Power Economics, and Deep Capture*, 152 *U. PA. L. REV.* 129 (2003) (exploring the significant implications of social psychology, especially situationism, for law and economics and law more generally).

32. See generally *HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT* (Thomas Gilovich et al. eds., 2002) (collecting essays about the various judgment biases that have been well documented in social psychology including the availability heuristic, correspondence bias, anchoring, optimism, *inter alia*).

33. See generally GARY S. BECKER, *THE ECONOMIC APPROACH TO HUMAN BEHAVIOR* (1976).

34. See generally Herbert Gintis, *Experimental Economics Will Foster a Renaissance of Economic Theory*, *J. ECON. BEHAV. & ORG.* (forthcoming), available at <http://www.umass.edu/preferen/gintis/VSmithReply.pdf>.

35. See DREW WESTIN, *THE POLITICAL BRAIN* 25–44 (2007).

socially constructed identity and who conforms his behavior accordingly; and who is heavily influenced by his situation, even while remaining blind to the situation and deeply attached to a set of rationales to explain his own behavior in dispositionist terms. That picture of human behavior goes some small measure toward explaining otherwise hidden social regulatory mechanisms; the question is whether they can be operationalized into law and policy.

Rather than dividing the following sections by sub-disciplines, each section involves a real-world scenario and a discussion of related empirical insights from the behavioral sciences. These insights are not comprehensive but rather are illustrative of the kinds of social regulatory processes that become visible—and therefore can be accounted for and harnessed in global regimes—with behavioral tools. Many of these behavioral insights have already been applied to areas of domestic law,<sup>36</sup> but none have been directly applied to international human rights law or policy—a task left for Part Three.

#### *A. Social Proof*

In the late nineteenth century, activists sought to end the Chinese practice of footbinding, a social custom that saw girls' feet mutilated to increase their aesthetic appeal.<sup>37</sup> The antfootbinding campaigns were a remarkable success: the entire practice came to a near halt less than a generation after the creation of the first antfootbinding societies. In their account of the campaign, political scientists Margaret Keck and Kathryn Sikkink say that “the most innovative technique” used by the activists was to have actors who agreed to fight footbinding announce their newly-reformed view and their avowal against the practice to their social community.<sup>38</sup> The authors note that “the members of antfootbinding societies pledged not to bind the feet of their daughters and to marry their sons only to women with unbound feet.”<sup>39</sup> This public avowal was a necessary step because it both helped to convey to uncertain actors information about how to behave, and it gave assurance to them that if they did not footbind, they would belong to and have access to a community of like-minded people.

Of course, it is difficult to isolate a singular mechanism to explain the rapid and near-total shift in Chinese footbinding behavior. It is possible that each family came to the conclusion independently through educational materials about the harms of footbinding. It is more plausible that education raised the prospect of a different social convention, and all that remained

---

36. See generally BEHAVIORAL LAW AND ECONOMICS (Cass Sunstein ed., 2000); ROBERT ELLICKSON, ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES (1991); Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998).

37. See generally WANG PING, *ACHING FOR BEAUTY: FOOTBINDING IN CHINA* (2000).

38. MARGARET E. KECK & KATHRYN SIKKINK, *ACTIVISTS BEYOND BORDERS* 63 (1998).

39. *Id.*

was attaining a critical mass who did not footbind.<sup>40</sup> But regardless of the mechanism, the mass decision to stop footbinding at one particular moment was likely at least partially the result of new information about peers' views of acceptable social behavior. Even if a majority of people privately did not support footbinding, they continued the practice and upheld a public-private distinction that created a situation of pluralistic ignorance.<sup>41</sup> In such situations—when a group of people privately rejects a norm but outwardly supports the norm thinking that their peers do as well—rapid shifts in behavior can occur with the mere introduction of information about actual peer preferences. Acting on this authority—the authority of one's peers—is a powerful phenomenon known in social psychology as social proof.<sup>42</sup>

In a series of littering experiments, psychologist Robert Cialdini has demonstrated that we often determine social convention by looking to what other people do. After putting flyers on peoples' cars and watching to see whether they littered, he found that people were significantly more likely to litter if there was already litter on the ground—which he suggested served as a signal about how people in that community handled unwanted pieces of paper.<sup>43</sup> Likewise, a study by Albert Bandura showed that children who had a phobia of dogs could be convinced to play with dogs after watching another child playing with a dog in a positive atmosphere for 10 minutes at a time over several days.<sup>44</sup> In another series of experiments using children, Bandura demonstrated that merely viewing aggressive behavior in adults led children to act more aggressively than those in the control group who had not had not been exposed to an aggressive model.<sup>45</sup> Perhaps the most compelling example of social proof comes from Solomon Asch, who found that subjects who could normally easily identify two lines of identical length had a very hard time doing so after a group of people in the same room had all judged the lines to be of different lengths.<sup>46</sup>

There are competing explanations for this phenomenon. Christina Bicchieri provides a rational explanation, noting: "The overriding motive, in this case, is the desire to imitate the choices of what one takes to be more

---

40. See, e.g., Gerry Mackie, *Ending Footbinding and Infibulation: A Convention Account*, 61 *AM. SOC. REV.* 999 (1996) (explaining the antifootbinding campaign's success in terms of its ability to make Chinese families aware of an alternative convention).

41. See TIMUR KURAN, *PRIVATE TRUTHS, PUBLIC LIES: THE SOCIAL CONSEQUENCES OF PREFERENCE FALSIFICATION* 78 (1995) (identifying "pluralistic ignorance" as the "misconceptions of preference distributions" that result when "people use mental shortcuts that violate basic statistical principles").

42. *Id.* at 162–67.

43. See Robert B. Cialdini et al., *A Focus Theory of Normative Conduct*, 24 *ADVANCES EXPERIMENTAL SOC. PSYCHOL.* 201 (1991).

44. See Albert Bandura et al., *Vicarious Extinction of Avoidance Behavior*, 5 *J. PERSONALITY & SOC. PSYCHOL.* 16 (1967).

45. See Albert Bandura et al., *Transmission of Aggressions Through Imitation of Aggressive Models*, 63 *J. ABNORMAL & SOC. PSYCHOL.* 575 (1961).

46. See SOLOMON E. ASCH, *SOCIAL PSYCHOLOGY* 451–59 (1952).

informed actors.”<sup>47</sup> Others suggest the mechanism is less cognitive and more automatic—occurring without any motive whatsoever.<sup>48</sup> Asch’s research, for example, suggests that the effects of social proof dissipated significantly if any one of the previous experimenters contradicted the group of erroneous statements—thereby opening an alternative “channel” for action.<sup>49</sup> Still, by acknowledging the phenomenon, psychologists have been able to identify some of the circumstances under which social proof can be harnessed for law and policy.

Some successful public health campaigns aimed at reducing binge drinking on college campuses, where the majority of students may be ignorant or mistaken about the preferences of their peers, inform their audiences that the majority of their peers do not like to binge drink, rather than inform them of the harms of binge drinking.<sup>50</sup> Similarly, successful campaigns to encourage energy-saving practices have relied on information about peer behavior.<sup>51</sup> The mere suggestion of how much energy was being used by neighboring homes was enough to drive energy use down in households that consumed more energy than the neighborhood average, and that shift in behavior was sticky enough to last even after the information about peer behavior was no longer available.<sup>52</sup> Social proof is perhaps the most compelling explanation for the sudden cascade of demonstrations in Egypt in 2008, which were organized through Facebook.<sup>53</sup> Not only did the social networking software give demonstrators a virtual space to assemble, but once available, Egyptians used the tool to reveal their private preferences to each other

47. BICCHIERI, *supra* note 20, at 80 (citing Christina Bicchieri & Yoshitaka Fukui, *The Great Illusion: Ignorance, Informational Cascades, and the Persistence of Unpopular Norms*, 9 BUS. ETHICS Q. 127 (1999)).

48. See, e.g., ROBERT CIALDINI, *INFLUENCE: SCIENCE AND PRACTICE* 3 (1985) (“Click and the appropriate tape is activated; whirr and out rolls the standard sequence of behaviors.”). See also *supra* text accompanying note 80.

49. “Channel factors,” originally identified by Kurt Lewin, have been shown to play a significant role in determining how people behave. Small first steps—minor sunk costs, a priori commitments, etc.—can lead to significant shifts in subsequent behavior. See, e.g., Marianne Bertrand et al., *Behavioral Economics and Marketing in Aid of Decision Making Among the Poor*, 25 J. PUB. POL’Y & MARKETING 8, 10 (2006) (“Another impressive illustration of a channel factor can be observed in Asch’s (1956) conformity studies, in which participants are led to make wildly misguided judgments that conform to those expressed by a group of the experimenter’s confederates. Remarkably, any dissent from unanimous opinion, even if it is in favor of a mistaken judgment, opens an appropriate channel, leading to an 80% reduction in participants’ tendency to conform.”).

50. See Christine M. Schroeder & Deborah A. Prentice, *Exposing Pluralistic Ignorance to Reduce Alcohol Use Among College Students*, 28 J. APPLIED SOC. PSYCHOL. 2150, 2174–75 (1998) (concluding that “pluralistic ignorance”—where mistaken views about the preferences of one’s peers lead to the hardening of a norm—plays a strong role in legitimizing binge drinking among college students). However, such campaigns can also cause boomerang effects. See, e.g., Robert Homik et al., *Effects of the National Youth Anti-Drug Media Campaign on Youths*, 98 AM. J. PUB. HEALTH 2229, 2235 (2008) (noting the risks of informing target actors of the behavior of their peers if those actors assume a lower baseline of behavior than the reality, in which case their behavior may “boomerang” up to meet the baseline rather than decrease).

51. See P. Wesley Schultz et al., *The Constructive, Destructive, and Reconstructive Power of Social Norms*, 18 PSYCHOL. SCI. 5, 429, 431 (2007).

52. *Id.* at 432.

53. See *supra* Introduction.

and to more accurately gauge whether or not they had the support of their peers in opposing governmental policies.

Social proof was also a key component of the propaganda strategy deployed in the Rwandan genocide, where radio broadcasts did not merely assert the purported reasons why Hutus should kill Tutsis but also attempted to draw on the community responsibility Hutus might feel toward one another, urging them to join their brothers who were already busy with the work of genocide.<sup>54</sup> The radio often sent cues—some more veiled than others<sup>55</sup>—to listeners about the behavior of their peers. Many radio broadcasts aimed at Hutus invoked a script suggesting that everyone else was already working to find and kill Tutsis and urging listeners to do their part.<sup>56</sup> These broadcasts served several purposes: to inform listeners of the behavior of others (other Hutus are killing Tutsis); to inform listeners of the prevailing norms (killing Tutsis is socially sanctioned); and to invoke shame and guilt on the part of those Hutus who were not hard at “work.”<sup>57</sup>

But if the strength of the information about peer behavior depends on the information available to the target audience, the effect of social proof will vary in different locations in a social network. And as the Rwandan example shows, social networks can be designed and manipulated to constrict or enable the flow of information.

### B. Social Networks

The Rwandan Genocide of 1994 was a carefully orchestrated social program that saw Hutu civilians destroy nearly a million Tutsis and Tutsi-sympathizers.<sup>58</sup> By most accounts, the genocide could not have happened without a network of radios to broadcast hate speech throughout the country. Human rights researcher Alison Des Forges reported that there was a coordinated program to arm all roadblocks and town centers with new radios that were considered genocidal tools as crucial for the killing as the

---

54. See ALISON DES FORGES, *LEAVE NONE TO TELL THE STORY* 258 (1999) (reporting on an early planning meeting among genocidaires, who “discussed the need for ‘uniformity and harmony’ of language”).

55. Many radio reports used euphemisms to pass messages to civilians about the killing transpiring around the country, without flagging unwanted attention abroad. See *id.* at 251 (“Many military officers understood that Bagosora and his supporters were saying one thing and doing another.”).

56. See *id.* (“Disseminating the message that ‘there is no place for moderates,’ RTLM heaped scorn on those who refused to participate: ‘The inhabitants of certain sectors don’t dare search! They say that the houses are occupied and that their owners are shut up inside them; they don’t dare search even in the banana groves!’”).

57. *Id.* at 272 (“Nzabonimana also addressed a large public meeting near the church of Kivumu, where ‘he asked the local population why they had not done their “work” and suggested that the Tutsi cattle were just waiting to be eaten.’”) (citations omitted).

58. See, e.g., *Rwanda: How the genocide happened*, BBC NEWS, Apr. 1, 2004, <http://news.bbc.co.uk/2/hi/africa/1288230.stm> (explaining how civilians contributed to the genocide, and estimating the death toll at 800,000). For more on the genocide as a carefully orchestrated social program, see DES FORGES, *supra* note 54 at 9, 15 (“preparing for the slaughter” and “popular participation”).

machetes.<sup>59</sup> Likewise, the radio was used to prime civilians to demand vengeance through violence immediately following the death of President Habyarimana. The radio broadcasts were brazenly transparent about their manipulation of the social fabric:

Call your brothers, all of you, board a boat and sail towards the mental Revolution Island. The island is none other than the CDR [the Hutu-power political party]. So now grab your oars, Hutus. Your disembarkment would no doubt be synonymous with vigilance and you will never again experience mental, administrative and economic domination.<sup>60</sup>

As field researcher Scott Straus later said about the media's roles in the genocide: "[T]he main effect of the radio broadcasts was to help establish killing Tutsis as the new order of the day—as the new 'law,' as the new basis for authority . . ."<sup>61</sup> By the time the genocidal violence was in full effect in the middle of 1994, many Rwandans had been cut off from independent sources of news, making them particularly dependant on Hutu-power media for information and social cues—and therefore particularly vulnerable to manipulation by hate speech.<sup>62</sup> Social network analysis helps to explain why.

Network science is an attempt to understand how "nodes"—individual units linked to other like units—interact based on their position in a given configuration of linked nodes; the ways in which the network is structured and how it changes; and the ways in which the individual nodes are affected by the network's features.<sup>63</sup> That work requires mapping the network—who are the players? What are their relationships?—and determining how they interact and how content moves between the various nodes. The central insight that network analysis offers is that information is relational, and that

59. See *id.* at 67 ("One foreign religious sister who traveled from Kibuye to Butare during the height of the genocide reported that she had seen new radios at every one of the dozens of barriers where she had been stopped en route.") (citation omitted).

60. Prosecutor v. Nahimana, Case No. ICTR-99-52-T, Judgment and Sentence, ¶ 916 (Dec. 3, 2003) ("In an article entitled 'Grab Your Oars Hutu,' signed by *Kangura* and published in May 1992 in *Kangura* No. International Version 10, the CDR was dubbed the 'mental Revolution Island' and Hutu readers were encouraged to join this revolution.")

61. Frank Chalk, *Intervening to Prevent Genocidal Violence: The Role of the Media*, in *THE MEDIA AND THE RWANDAN GENOCIDE* 375, 376 (Allan Thompson ed., 2004) (quoting Scott Straus, *The Order of Genocide: Race, Power, and War in Rwanda* (2004) (unpublished Ph.D. dissertation, University of California, Berkeley)).

62. See DES FORGES, *supra* note 54, at 369 ("The restrictions on movement and the cancelling of the Friday market meant they could not gather news from others as they usually did.")

63. See Katherine J. Strandburg et al., *Law and the Science of Networks: An Overview and an Application to the "Patent Explosion,"* 21 BERKELEY TECH. L.J. 1293, 1300–01 (2006) (defining the goal of network science as threefold: "(1) to measure, describe, and categorize network structure and the patterns of relationships between network nodes; (2) to understand network evolution and growth and its relationship to network structure; and (3) to understand how the collective behavior of entities connected in a network depends on and derives from the network's structure.")

meta-data about relationship structures can offer insight into social dynamics.

Networks occur naturally, but as we see in the Rwanda example, they can be consciously constructed.<sup>64</sup> Social engineers may seek to enhance or limit the extent to which nodes can communicate with one another—that is, redesign the structure of the network—based on their objectives. An example of this is the recent introduction of wireless communications in the helmets of American football defensive lines, which allowed their network of coaches to communicate with them from the sidelines while explicitly leaving opposing teams out of the communication network whereas previous hand-signal networks had been susceptible to interference.<sup>65</sup>

Networks can also be reconfigured to enhance the distribution of norms, and access can be strategically expanded or curtailed.<sup>66</sup> For the Rwandan genocidaires, that meant introducing rebel leaders to each other in early planning meetings and later arming all check points with radios—ensuring that they acted as physical as well as informational nodes in the genocidaire network.<sup>67</sup> Furthermore, they cut civilian access to outside information and replaced information networks with Radio RTLM broadcasts, upon which many civilians were dependent for information, situating them in a position of network weakness.<sup>68</sup>

Many networks remain robust until the central nodes are wiped out. In the Rwandan informational network, broadcast radios connected listeners to the “hub” nodes of the broadcasters, but the radios were one-way transmissions, making the listeners dependent nodes. Meanwhile, the civil war disrupted normal channels of information and communication, cutting the listeners out of other informational networks, making them that much more dependent on the radio and increasing the central hub’s power.

64. An example that occurs in nature is cricket chirping across a field, which is distributed in a network pattern similar to that of a communicable disease. See DUNCAN WATTS, *SMALL WORLDS* xiii (1999). The most prominent man-made network to date is, of course, the Internet.

65. See Jim Corbett, *Can You Hear Me Now? Radio Calls in NFL Defenders’ Helmets*, USA TODAY, Aug. 27, 2008, at 1C, available at [http://www.usatoday.com/sports/football/nfl/2008-08-27-helmet-radios\\_N.htm](http://www.usatoday.com/sports/football/nfl/2008-08-27-helmet-radios_N.htm).

66. See generally David A. Lake & Wendy H. Wong, *The Politics of Networks: Interests, Power, and Human Rights Norms*, in *NETWORKED POLITICS: AGENCY, POWER, AND GOVERNANCE* (Miles Kaher ed., 2009) (using network analysis to explain the growth of advocacy group Amnesty International and its ability to manipulate the size and scope of a network).

67. See DES FORGES, *supra* note 54, at 67 (“The radio was to become even more effective in delivering the message of hate directly and simultaneously to a wide audience. Before the war, Rwanda had only one radio station, the national Radio Rwanda, but listening to the radio was a popular distraction among ordinary people and elite alike. In 1991, some 29 percent of all households had a radio. The number of radio sets was presumably much higher by the start of the genocide. In some areas, the government distributed radios free to local authorities before the genocide and they may have done so after the killing began as well.”) (citations omitted).

68. See *id.* at 68 (“This national radio sometimes broadcast false information, particularly about the progress of the war, but most people did not have access to independent sources of information to verify its claims.”).

These insights help to explain some of the social power that Hutu genocidaires had at their disposal; they can also be used by courts to understand the genocidal mechanisms after the fact.<sup>69</sup> But ultimately the network is only as powerful as the information it transmits.

### C. Social Meaning

Countries like Ghana are plagued by diseases that could easily be eliminated through the regular practice of handwashing with soap after toilet use.<sup>70</sup> Ghanaians regularly wash their hands after toilet use, but as a population they rarely use soap despite the wide availability of it.<sup>71</sup> In 2000, Val Curtis, the director of the Hygiene Center at the London School of Hygiene and Tropical Medicine, was fed up with trying to convince people of the medical necessity of soap use; she had been explaining to people how soap kills germs for years, to little effect. She helped to form an initiative called the “Global Public Private Partnership for Handwashing with Soap,” and convinced some of the world’s largest consumer products manufacturers to participate.<sup>72</sup>

Curtis reached out to Unilever—one of the many companies that have invested heavily in linking habits with their products—which set out to study the social situation of soap use in Ghana. Ghanaians, it turned out, regularly washed their hands after going to the bathroom, but usually without soap. The only time they used soap was when they felt disgust—for example after handling grease—and surveys showed that they viewed going to the bathroom as natural and not disgusting.<sup>73</sup> So Curtis, with the help of Unilever’s marketing team, designed a series of advertisements designed to make people feel disgust in bathrooms. One of the advertisements showed a red spot on someone’s hands after they used the bathroom, and the spot followed them and could be seen on everything they touched.<sup>74</sup> Only washing with soap made the red stain go away. An infectious-disease specialist who worked with Curtis on the Ghana campaign explained that, “This was radically different from most public health campaigns. . . . There was no mention of sickness. It just mentions the yuck factor. We learned how to do that from the marketing companies.”<sup>75</sup>

---

69. See *infra* Part III.B.

70. See Valerie A. Curtis et al., *Masters of Marketing: Bringing Private Sector Skills to Public Health Partnerships*, 97 AM. J. PUB. HEALTH 634, 634–35 (2007); Charles Duhigg, *Warning: Habits May Be Good for You*, N.Y. TIMES, July 13, 2008, at BU6.

71. Curtis et al., *supra* note 70, at 636; Duhigg, *supra* note 70.

72. Curtis et al., *supra* note 70.

73. See Curtis et al., *supra* note 70, at 636–37; Duhigg, *supra* note 70.

74. Curtis et al., *supra* note 70, at 636; Duhigg, *supra* note 70.

75. Duhigg, *supra* note 70.



The advertisements seem to have worked: Curtis reported an increase in handwashing with soap after toilet use by thirteen percent and an increase before eating of forty-one percent.<sup>76</sup> But why did they work?

One explanation is that the advertisements changed the social meaning of handwashing with soap. Once the meaning of soap use was altered, it triggered an emotional response that brought with it the desired behavior. The campaign made use of what are known as “implicit attitudes,” which engender automatic responses. A growing body of literature shows that our attitudes and preferences are often implicit, instantaneous, and hidden (i.e., unconscious), even if we have good reasons for them.<sup>77</sup> Indeed, many of these reasons are arrived at after the implicit attitude is formed as an *ex post* justification. For example, the preference for in-group members over out-group members has been shown to be automatic.<sup>78</sup> In a famous series of studies, experimenters showed that stereotypical knowledge is activated automatically—e.g. “black” and “athlete” are associated automatically, while “white” and “good” are associated automatically.<sup>79</sup>

Cialdini calls such automatic influences “click-whirr” heuristics because typically once the trigger or “click” occurs, the response is as automatic as the whirr of a machine clicking into gear.<sup>80</sup> Studies have shown that automatic influence is manipulable: “[E]xposure to unfamiliar or counter-stereotypical associations can temporarily reshape implicit attitudes; with the right situational manipulations, for example, subjects begin to automatically associate the terms ‘strong’ and ‘female.’”<sup>81</sup> When we think we have come to a carefully reasoned conclusion, we may in fact be justifying the natural and automatic response to a particular call.

#### D. Social Roles

In the American South in the mid-1990s, some churches began urging their teenage members to take “virginity pledges”: oaths that they would not have sexual intercourse before marriage. The pledges were widespread<sup>82</sup>

76. *Id.*; Curtis et al., *supra* note 70, at 637.

77. See, e.g., Mahzarin R. Banaji, *Implicit Attitudes Can Be Measured*, in *THE NATURE OF REMEMBERING: ESSAYS IN HONOR OF ROBERT G. CROWDER* 117, 118 (H.L. Roediger, III et al. eds., 2001) (noting that psychologists have recently discovered and begun to measure “new, previously undetected forms of attitudes,” specifically those that are implicit and unconscious).

78. See Jon Hanson & David Yosifon, *The Situational Character: A Critical Realist Perspective on the Human Animal*, 93 *GEO. L.J.* 1, 89 (2004) (discussing experiments that show that judgments about other sexes and races are unconscious).

79. *Id.* See also Jason Mitchell, Brian Nosek & Mahzarin Banaji, *Category Salience Determines Implicit Attitudes Toward Black Female and White Male Targets* (Paper presented at the First Annual Meeting of the Society for Personality and Social Psychology, Nashville, Tenn., 2000), available at <http://projectimplicit.net/nosek/talk/SPSP.mitchell.handout.2000.doc>.

80. See CIALDINI, *supra* note 48.

81. Hanson & Yosifon, *supra* note 78.

82. The “True Love Waits” program alone claims over two million pledgers. See *True Love Waits Launches Community-Wide Initiative*, BAPTIST PRESS, Feb. 14, 2005, <http://www.bpnews.net/bpnews.asp?ID=20143>.

and have been shown to have had a mitigating impact on the sexual activity level of teens who took the oath.<sup>83</sup> Not all teens who took the pledge refrained from sexual activity before marriage—a significant number broke their pledge. When asked whether they had broken their pledge, a full seventy-three percent of surveyed pledge-breakers (individuals who took the pledge and reported becoming newly sexually active thereafter) denied having taken a pledge in the first place.<sup>84</sup> Such a discrepancy could indicate poor memory or, as one prominent study suggested, it may indicate that the survey subjects seek to see themselves as consistent with their word.<sup>85</sup>

We tell stories to ourselves and others, and these stories construct our identity, which in turn plays a role in determining our behavior.<sup>86</sup> We bend our behavior and our justification for that behavior according to our narrative-identity. For example, even when we make commitments for purely financial reasons, those commitments can last long after the financial interest is gone. Such commitments, we say, have “grown their own legs”—they have been incorporated into our own identity, and we will continue to stay consistent with them,<sup>87</sup> much the way we aim to hold consistent to our written commitments.<sup>88</sup>

This behavior is known as the commitment/consistency bias, which holds that once we make a commitment we will go to great lengths to conform our behavior to that commitment, even ignoring or not seeing our deviations from it in order to maintain a consistent view of ourselves and our behavior.<sup>89</sup> This feature can be manipulated to induce behavioral change. For example, Edgar Schein has shown how complex psychological operations in Chinese prisons during the Korean War led to compliance by American prisoners-of-war (“POWs”).<sup>90</sup> Commitments by the American POWs as

83. See Hannah Brückner & Peter Bearman, *After the Promise: The STD Consequences of Adolescent Virginity Pledges*, 36 J. ADOLESCENT HEALTH 271 (2005) (noting that while sexual encounters and therefore risk factors were reduced among pledgers, sexual transmission rates were as high as those among non-pledgers, potentially indicating a lower use of condoms among pledgers at their sexual debut).

84. Janet Rosenbaum, *Reborn a Virgin: Adolescents' Retracting of Virginity Pledges and Sexual Histories*, 96 AM. J. PUB. HEALTH 1098, 1101 (2006) (noting that in some cases “self-identified virginity pledgers chose to have sexual intercourse despite their pledge and may later have ‘overlooked’ the pledge as a result of cognitive dissonance between their virginity pledge and their history of sexual intercourse.”).

85. See *id.*

86. See generally ERVING GOFFMANN, *THE PRESENTATION OF SELF IN EVERYDAY LIFE* (1956).

87. See CIALDINI, *supra* note 48, at 92–103 (showing that people who are low-balled into saving home energy through a promise to run their names in a newspaper congratulating them if they conform will continue with their commitment even after the reward is removed. The commitments “grow their own legs”—that is, if they initially reduced energy use for status or to conform with social norms, the modified behavior remains because the subjects’ self conception has changed. (citing Michael Pallak et al., *Commitment and Energy Conservation*, in APPLIED SOC. PSYCHOL. ANN. (L. Bickman ed., 1980))).

88. See CIALDINI, *supra* note 48, at 75–81 (describing the effects of written commitments).

89. See CIALDINI, *supra* note 48, at 57.

90. See Edgar H. Schein, *Some Observations on Chinese Methods of Handling Prisoners of War*, 20 PUB. OPINION Q. 321 (1956).

small as written statements—even those made facetiously—eventually unsettled prisoners' convictions against Chinese communism.<sup>91</sup>

Public commitments are even more binding than private ones, and as Richard Katzev and Theodore Wang have shown, written commitments are stronger than oral ones.<sup>92</sup> Written commitments can lead to pro-social behavioral changes, ranging from recycling, electricity use, and bus ridership—even without the initial corollary attitude change, meaning that the subject perhaps never intended to follow the initial commitments.<sup>93</sup> One explanation for the strength of the commitment/consistency bias is the strength of the narrative about our identity, which carries a set of norms with it.<sup>94</sup>

Along with our narratives about our individual and group identities, we create narratives to explain or justify who gets included in a group and who is excluded. These narratives drive the phenomenon of in-group and out-group bias, where group members favor group members and disfavor non-group members. These biases can be harnessed and manipulated for pro-social ends: in-group biases can be productive when they increase cooperation and in some cases allow groups to overcome prisoner's dilemma problems.<sup>95</sup> Likewise, discussion raises in-group cooperation—perhaps because it allows time for group identity to be formed.<sup>96</sup> Interestingly, these aims can be achieved without significant buy-in (psychological, social, or cultural) from the group members.<sup>97</sup> This construction can happen quite artificially. As Jon Hanson notes:

Merely telling [people] that they are now a group leads them to want to reward their own group more, and to see its members as having better personalities, nicer looks, less responsibility for any failures, and more responsibility for successes. The self-serving attributional biases of individuals also appear on the group level as in-group members [tend] to attribute internal [or dispositional]

---

91. There is, however, a question about the lasting effect of these statements. Many of the soldiers returned to their previous views after leaving the confines of the camp. *See id.* at 326.

92. *See* Richard Katzev & Theodore Wang, *Can Commitment Change Behavior? A Case Study of Environmental Actions*, 9 J. SOC. BEHAV. & PERSONALITY 13, 22 (1994). *See also* Theodore H. Wang & Richard D. Katzev, *Group Commitment and Resource Conservation: Two Field Experiments on Promoting Recycling*, 20 J. APPLIED SOC. PSYCHOL. 265 (2006) (finding that group commitments can be lasting, but only with monitoring, while individual commitments lasted beyond the trial period).

93. *See* CIALDINI, *supra* note 48, at 94.

94. *See* BICCHIERI, *supra* note 20, at 94 (noting that “[s]ocial norms are embedded into scripts.”).

95. *See id.* at 146–47 (citing Dawes et al., *Not Me or Three But We: The Importance of Group Identity in Eliciting Cooperation in Dilemma Situations*, 68 ACTA PSYCHOLOGICA 83, 83–97) (noting that Dawes et al. “reasoned that if individuals incorporate groups into their self-concept, a motivational shift would occur and group welfare would matter more than individual welfare”).

96. *Id.* at 147.

97. *See id.* at 158 (noting that “[i]t is the realization that there can be psychological group membership without interdependence, need satisfaction, personal attraction, social structure, or common norms and values that led Tajfel, and later Turner and Brewer, to design experiments in the context of the minimal group paradigm”).

causes to positive in-group and negative out-group behavior and to attribute negative in-group behavior and positive out-group behavior to external [or situational] causes.<sup>98</sup>

This helps to explain the persistence of stereotypes.<sup>99</sup>

Describing what they call “cognitive groupism,” Hanson and Yosifon analyze Muzafer Sherif’s famous experiment in which he ran a summer camp, with young boys as subjects divided into two groups, called the “Eagles” and “Rattlers.”<sup>100</sup> Throughout the summer, Sherif measured the boys’ relations, and found that when pitted against each other in direct competitions, even for games, they took on strong in-group effects and heavy out-group prejudices.<sup>101</sup> The group identities and related biases were so strong they were difficult to ramp down with information alone.<sup>102</sup>

It was not until the boys had to cooperate together under a situation of shared challenge—fixing their bus, which they had been told was broken down—that they described members of the other group in positive terms, suggesting that the situation of cooperation forced them to shed their in-group biases.<sup>103</sup> These biases are not unique to children or summer camps. Groups and subsequent in-group biases can be created through such artificial distinctions as grouping people by a randomly assigned number.<sup>104</sup>

### E. Other Biases

Social association also determines the way we process new information. As Dan Kahan and his co-authors have shown, the so-called “Factual Enlightenment Model” of human decisionmaking—upon which much of the law is based—rests on two faulty assumptions: that humans prefer truth to falsity; and that social interactions transmit true statements.<sup>105</sup> In fact, they say, “culture is prior to facts in individual cognition,” and that “[t]hrough an overlapping set of psychological and social mechanisms, individuals adopt the factual beliefs that are dominant among persons who share their cultural

98. Hanson & Yosifon, *supra* note 78, at 100 (citations omitted).

99. See Ronald Chen & Jon Hanson, *Categorically Biased: The Influence of Knowledge Structures on Law and Legal Theory*, 77 S. CAL. L. REV. 1103, 1184–85 (2004) (citing John Duckitt, *Psychology and Prejudice: A Historical Analysis and Integrative Framework*, 47 AM. PSYCHOLOGIST 1182, 1188–89 (1992) and Henri Tajfel & John C. Turner, *The Social Identity Theory of Intergroup Behavior*, in THE PSYCHOLOGY OF INTERGROUP RELATIONS 7, 11 (Stephan Worchel & William G. Austin eds., 2d ed. 1986)). See also Steven Fein & Steven J. Spencer, *Prejudice as Self-Image Maintenance: Affirming the Self Through Derogating Others*, 73 J. PERSONALITY & SOC. PSYCHOL. 31, 31 (1997).

100. Hanson & Yosifon, *supra* note 78, at 55.

101. MUZAFER SHERIF ET AL., INTERGROUP CONFLICT AND COOPERATION: THE ROBBERS CAVE EXPERIMENT 27–68 (1961).

102. See Hanson & Yosifon, *supra* note 78, at 54–58.

103. SHERIF ET AL., *supra* note 101, at 171–78.

104. See Henri Tajfel, *Experiments in Intergroup Discrimination*, 223 SCI. AM., 96–102 (1970) (finding that after playing a simple computer game that resulted in a numerical score, boys favored those who had received a similar score and discriminated against those with different scores).

105. See Dan M. Kahan et al., *Modeling Facts, Culture, and Cognition in the Gun Debate*, 18 SOC. JUST. RES. 283, 286 (2005).

orientations.”<sup>106</sup> In a series of simulations, Kahan et al. show that cultural cognition—awareness of what culturally similar people think or do—can crowd out empirical data that might otherwise change the subject’s mind.<sup>107</sup>

In a related study, the authors found proof for political scientist Aaron Wildavsky’s claim that “culture” or social relations are the unifying heuristic for all other heuristics.<sup>108</sup> In an original national survey (N=1843), they found that Wildavsky-style cultural orientations accounted for policy attitudes on gun control, environment, capital punishment, and gay marriage, after controls.<sup>109</sup> Their claim is not that culture determines peoples’ views per se, but rather that culture orients people to process new information in a particular way: “Culture conflict is not the object of democratic politics; it is a byproduct of the social and psychological processes through which culturally diverse citizens *apprehend* their common interests.”<sup>110</sup> The result is that policy issues can be framed in such a way so as to ensure they are processed through an essentially cultural medium, and therefore crowd out more rational results. In a sense, this means that culture can be “hijacked” for anti-social ends.

These are but a few of the powerful examples of biases that influence individual and group behavior. A great number of other known biases explain how and why human thought and behavior are influenced by social situations. These include: affect,<sup>111</sup> biased assimilation,<sup>112</sup> the availability heuristic,<sup>113</sup> the endowment effect,<sup>114</sup> shortsightedness,<sup>115</sup> system justifica-

106. *Id.* at 285.

107. *See id.* at 292–97 (designing and simulating what happens when two competing models for information processing exist simultaneously, what the authors call the “truth vs. culture” model).

108. *See* John Gastil et al., *The “Wildavsky Heuristic”: The Cultural Orientation of Mass Political Opinion* (Yale Law Sch., Pub. L. Working Paper No. 107, 2005), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=834264](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=834264).

109. *Id.* at 2.

110. *Id.* at 31.

111. The gut feeling of liking, or affect, one has at the time of a decision can dominate and override so-called rational decisionmaking processes. This was demonstrated in a series of experiments in which subjects were shown often nonsensical objects or words and asked to say which things they liked more than others, or which seemed more familiar, or both. *See* R. B. Zajonc, *Feeling and Thinking: Preferences Need No Inferences*, 35 *AM. PSYCHOLOGIST* 151 (1980).

112. *See* Kahan et al., *supra* note 105, at 292 (“Unlike rational Bayesian information-processors, individuals don’t update their prior beliefs based on new evidence; instead they evaluate new evidence based on its conformity to their priors, dismissing as unpersuasive evidence that contradicts their existing beliefs.”).

113. *See generally* Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 5 *COGNITIVE PSYCHOL.* 207, 211–12 (1973) (asking people to estimate the percentage of words that begin with “R” or “K” as opposed to words having those letters in the third position, and finding that presumably because the subjects could easily recall examples in which words began with those letters, but did not have available in their minds examples of words with those letters in the third position, subjects overestimated the proportion of words beginning with the letters “R” or “K”).

114. *See* Richard Thaler, *Toward a Positive Theory of Consumer Choice*, 1 *J. ECON. BEHAV. & ORG.* 39, 43–47 (1980) (noting the fact that people tend to demand much more to give up an object than they would pay to obtain it). This has particular significance for battles over scarce resources. *See, e.g.*, Ziv Caromon & Dan Ariely, *Focusing on the Forgone: How Value Can Appear So Different to Buyers and Sellers*, 27

tion,<sup>116</sup> and framing effects.<sup>117</sup> What emerges from a brief examination of these biases is the fact that the circumstances that influence individual behavior are much more powerful than we think. And yet, as the next Part shows, very little of this learning has been incorporated into formal human rights practice or scholarship.

## II. IGNORING HUMAN BEHAVIOR

There is widespread skepticism about the efficacy of international human rights treaties.<sup>118</sup> In order to determine the strength of human rights law as a regulatory regime, many scholars look solely to whether states' behavior is limited by their treaty agreements. Framing the issue in this way leads some critics to conclude that international human rights law is hopelessly weak because, by design, it has no central enforcement mechanism.<sup>119</sup> Meanwhile, human rights practitioners seek stronger enforcement mechanisms in order to make human rights law more like domestic law.<sup>120</sup>

---

J. CONSUMER RES. 360, 362 (2000) (finding holders of scarce basketball tickets will sell for an average price of \$2411 whereas buyers will pay an average of \$166).

115. See Samuel M. McClure et al., *Separate Neural Systems Value Immediate and Delayed Monetary Rewards*, 306 SCI. 503 (2004), cited in JON ELSTER, *EXPLAINING SOCIAL BEHAVIOR* 118, 123 (2007) (noting that humans tend to value welfare and benefits in the present more than in the future).

116. Also known as the status quo bias, system justification is the tendency to try to see the status quo as fair, legitimate, and good. This phenomenon would be expected in dominant groups, but has been observed even in people and groups in a non-dominant or minority position. See John T. Jost et al., *A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo*, 25 POL. PSYCHOL. 881 (2004) (synthesizing many previous experiments and hypotheses about system justification, with special focus on "implicit outgroup favoritism among members of disadvantaged groups"). This is especially true of people in positions of uncertainty, transition, or danger: "[p]eople who possess heightened needs to manage uncertainty and threat are especially likely to embrace conservative, system-justifying ideologies. . . ." John T. Jost & Orsolya Hunyady, *Antecedents and Consequences of System-Justifying Ideologies*, 14 CURRENT DIRECTIONS IN PSYCHOL. SCI. 260, 261 (2005).

117. Framing effects—that is, drawing different conclusions about data based on how they are presented—have been well documented in psychology and social movement literatures. See Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCI. 453 (1981) (noting that decisions were made differently depending on the way the decision was framed). See also Marshall Ganz, *The Power of Story in Social Movements* (Aug. 2001) (unpublished manuscript, on file with Harvard University) (recounting Cesar Chavez's narrative vision for the United Farm Workers social movement—linking narrative production to identity through leadership). This has even been shown in fMRI brain scans. See Benedetto De Martino et al., *Frames, Biases, and Rational Decision-Making in the Human Brain*, 313 SCI. 84 (2006) (showing fMRI scans of amygdala and orbital and media prefrontal cortex regions of the brain as subjects make financial decisions that are framed in the same way as the 1981 Tversky and Kahnemann study).

118. See JACK L. GOLDSMITH & ERIC A. POSNER, *THE LIMITS OF INTERNATIONAL LAW* 108 (2005) (arguing that "modern multilateral human rights treaties have little exogenous influence on state behavior").

119. See *id.* at 120 ("Nor do the modern human rights treaties have an effective or reliable coercive enforcement mechanism.").

120. See, e.g., Desmond Tutu & Mary Robinson, *Rights Demand Strong Institutions: Sixty Years After the Universal Declaration of Human Rights, We Need Legislative Muscle to Back Up Our Ideals*, THE GUARDIAN (London), Dec. 7, 2008, <http://www.guardian.co.uk/commentisfree/2008/dec/07/unitednations-humanrights> ("[B]illions of people are today unable to access or protect their legal rights because judicial and

This is ironic because domestic legal theory has witnessed a significant trend away from top-down law enforcement models toward those that harness community social structures to enforce community norms.<sup>121</sup> Where the previous law enforcement models depended on police authority to enforce laws, many newer approaches rely at least partially on the strength of a particular social situation to trigger the desired norm that may in turn regulate the target behavior.<sup>122</sup> This trend in domestic law—away from top-down police enforcement and toward situation-based community regulation—may offer a useful model for human rights scholars and regime designers. This Part shows how both international legal scholars and human rights regime designers are, in their concern with international regulation of state acts or state regulation of individual acts, insufficiently attentive to empirical evidence about the social situations that influence those acts.

#### A. *International Human Rights Scholarship*

Attempts to determine the robustness of international human rights law abound, but they under-explain the mechanism through which human rights laws affect ground-level behavior. While the methodologies vary widely, most empirically-grounded theories about human rights remain an attempt to answer the question “Why Do Nations Obey International Law?”<sup>123</sup> These state-centric theories under-appreciate alternative avenues for regulatory influence. Specifically, they under-appreciate the power of local social conditions over the violation and/or protection of human rights.

The behavioral approach proposes a different starting point: rather than assuming an international norm and asking how that norm can be enforced as a rule on states, it asks how international human rights regimes could pay more attention to the social situations that give or deny its norms social meaning. To begin that task, one must make an inquiry into the nature of norms, social situations, and human behavior. But current legal scholarship about human rights is generally unconcerned with social science insights about human behavior.<sup>124</sup> Instead, scholars have focused on the question of

---

law enforcement systems are impoverished or lack integrity. Changing this will require massive investment in courts, judicial officials, police, prison systems, social ministries, and parliaments.”).

121. See Kahan & Meares, *supra* note 22. See also Lanni, *supra* note 22.

122. These community justice models are not entirely proven as policing techniques. But without entering a debate about their efficacy—or the inefficacy of earlier approaches—it is enough for present purposes to note that they offer the prospect of viewing an old problem in a new light.

123. See Harold Hongju Koh, *Why Do Nations Obey International Law?*, 106 *YALE L.J.* 2599 (1997).

124. One exception is the scholarship on conformity in mass atrocity. See, e.g., MARK A. DRUMBL, *ATROCITY, PUNISHMENT, AND INTERNATIONAL LAW* 29–35 (2007) (discussing empirical evidence from social psychology that seeks to explain how individuals conform with the group to commit atrocities). But the author’s search of Westlaw database “Journals and Law Reviews” with the following query produced zero relevant results: (“fundamental attribution error” /p “human rights”). Similar searches were equally unfruitful. Searching for (“fundamental attribution error” & “international law”) in the same database produced only 25 articles, none of which provided an in-depth analysis of the relationship between the two terms. These searches were last done October 29, 2009.

why states follow international law, with most answers relying on empirical claims about state behavior at the international level.

The recent debate over what Jack Goldsmith and Eric Posner call *The Limits of International Law* is instructive.<sup>125</sup> In that book, the authors resurrect a realist claim that international human rights law does not significantly limit state behavior and therefore does not have the effect its proponents claim. Contrasting with this realist account are many constructivist tales about the dynamic relationship between states and the international community.<sup>126</sup> As political scientist Alistair Iain Johnston notes, these scholars differ from realists in their shared assumption that “social interaction in international relations can affect actor interests in such a way as to then change the fundamental characteristics of the normative structures that constitute the world political system.”<sup>127</sup> This constructivist account is behavioral in the sense that it is concerned with the ‘social’ situation of states in the international order; but these state-based inquiries are often more concerned about law enforcement by states even where that law enforcement does not completely bring with it the attendant norm. Despite early work by constructivists to explain domestic social dynamics and human rights norm internalization,<sup>128</sup> recent constructivist accounts of international law that explicitly account for social dynamics often pay little or no attention to the ground level situations inhabited by the people human rights law is meant to address.

For example, legal scholars Ryan Goodman and Derek Jinks offer a novel model whereby states become acculturated to their international situation and mimic certain attributes of their peers.<sup>129</sup> They claim that “[a]cculturation induces behavioral changes not only by changing the target actor’s incentive structure or mind but also by changing the actor’s social environment.”<sup>130</sup> The mechanism functions through cognitive and social pressures—real or imagined—applied by a group, such as shaming, shun-

125. See GOLDSMITH & POSNER, *supra* note 118, at 134. The book sparked a debate in the international legal community that has been covered well elsewhere, so we will not revisit it here other than to note the narrow scope of the debate over the reach of international human rights law. See Oona A. Hathaway & Ariel N. Lavinbuk, *Rationalism and Revisionism in International Law*, 119 HARV. L. REV. 1404 (2006) (critically reviewing *THE LIMITS OF INTERNATIONAL LAW*).

126. See Gerard Ruggie, *What Makes the World Hang Together? Neo-Utilitarianism and the Social Constructivist Challenge*, 52 INT’L ORG. 855 (1998) (explaining the roots and development of constructivism and making the distinction between three forms of constructivism: neo-classical constructivism, post-modern constructivism, and naturalistic constructivism). Ruggie notes that “[t]he first instinct of the willing constructivist is to incorporate norms, identities, and meaning into the study of international relations with minimum disruption to the field’s prevailing epistemological stance, on which hopes for analytical rigor and cumulative knowledge are believed to rest.” *Id.* at 884.

127. ALASTAIR IAIN JOHNSTON, *SOCIAL STATES: CHINA IN INTERNATIONAL INSTITUTIONS, 1980–2000*, at xiv (2007).

128. See, e.g., *THE POWER OF HUMAN RIGHTS* (Thomas Risse, Stephen C. Ropp & Kathryn Sikkink, eds., 1999).

129. See Ryan Goodman & Derek Jinks, *How to Influence States: Socialization and International Human Rights Law*, 54 DUKE L.J. 621, 626 (2004).

130. *Id.* at 638.



ning, and back patting.<sup>131</sup> While the model draws on social science insights about human behavior, scholars have chafed at the suggestion that states can be understood through behavioral insights about human behavior.<sup>132</sup>

Despite these critiques, the mimicry mechanism proposed by Goodman and Jinks has already been applied to general theories of states' ratification of and compliance with treaties, most recently by Michael Stein and Alex Geisinger.<sup>133</sup> But those authors, like Goodman and Jinks, Hathaway, and Goldsmith and Posner, focus on treaty adoption and enforcement as a metric for international law, which means their work is ultimately: a) fundamentally concerned with the state as the relevant actor, rather than the individual's ground-level social situation; and b) legalistic in its fetishism of rule-ratification and rule-enforcement. This raises the question of how the norms behind international legal obligations are meant to change individual behavior in the domestic realm.

Writing in response to Goodman and Jinks, Harold Koh notes:

Goodman and Jinks also have yet to develop the details of their discussion of *domestic* channels of influence, explaining, for example, how a society can move from a textual internalization of international norms . . . to "internalization on the ground," in the sense of genuine social acceptance of human rights norms (what I call elsewhere the process of "social internalization"). Their model also unduly emphasizes positive treaty law over unwritten customary law and makes insufficient use of the notion of an "interpretive community," in which the meaning of global norms can be debated and refined through the interaction of external and internal interpreters.<sup>134</sup>

Understanding state behavior is crucial to designing international legal regimes, but it is not the only way to approach the question of how, or whether, international law has meaning.<sup>135</sup> In the case of human rights law, which makes individuals the subject of international law, the focus on state behavior has distracted scholars from the question of how normative influ-

131. *Id.* at 641.

132. See José Alvarez, *Do States Socialize?*, 54 DUKE L.J. 961, 969 (2005) (responding to Goodman and Jinks by characterizing their anthropomorphic account of state behavior as "pop psychology without people").

133. See Alex Geisinger & Michael Ashley Stein, *A Theory of Expressive International Law*, 60 VAND. L. REV. 77 (2007) (developing an expressive theory under which states are motivated to develop their norms and to ratify and comply with international law by their "desire for esteem from the global community").

134. Harold Hongju Koh, *Internalization Through Socialization*, 54 DUKE L.J. 975, 981–82 (2005) (citations omitted).

135. See generally Paul Schiff Berman, *From International Law to Law and Globalization*, 43 COLUM. J. TRANSNAT'L L. 485 (2005) (calling for a new field—"law and globalization"—which has its theoretical grounding much more in *jus gentium* than the *jus inter gentes* that grounds current statist approaches to international law).

ence works in the social situations where rights are given or deprived of meaning. Even if new data conclusively resolved the above debates—perhaps the liveliest in international human rights law in the past decade of American scholarship—it will have contributed little to our understanding of how international human rights law affects, or can affect, behavior “on the ground.” Nor does the debate help us to understand the role that nonstate actors play in effecting international legal change. In many places where recalcitrant states are reluctant to enforce international commitments at home, those nonstate actors may necessarily have a bigger role or perhaps the only role in promoting, or indeed creating, the situations that enable human rights norms.

Law and economics scholars long ago identified social norms as a supra-legal means of regulating behavior.<sup>136</sup> Norms have been used as explanatory tools for a whole host of domestic regulatory regimes, from crime,<sup>137</sup> to cattle ranching,<sup>138</sup> to seat belt use.<sup>139</sup> Social norms have also been applied to international law as an explanation for how and why states ratify treaties, and how those treaties then become internalized through domestic channels.<sup>140</sup> But despite the plethora of human rights scholarship declaring the need for tools to understand and build “human rights cultures,”<sup>141</sup> there is surprisingly little work applying the regime-design lessons from the rich social norms literature to the challenge of promoting international human rights norms. This failure is striking because human rights norms have as much potential to regulate individual behavior as any other set of norms, and possibly much more than most.<sup>142</sup>

Legal scholars have not all been blind to ground level situations where international norms are translated into local communities.<sup>143</sup> Koh popularized a theoretical groundwork for human rights norm diffusion with his theory of “transnational legal process.”<sup>144</sup> That model explicitly relies on the success of the domestic social norms literature. Koh explains that “[w]e intuitively understand that the best way to get people to worship, to stop

---

136. See, e.g., POSNER, *supra* note 24, at 11–35.

137. See Kahan and Meares, *supra* note 22.

138. See ELLICKSON, *supra* note 36.

139. See LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 93–94 (1999).

140. See Harold Hongju Koh, *Transnational Legal Process*, 75 NEB. L. REV. 181 (1996).

141. This literature is vast. See generally CULTURAL ANALYSIS, CULTURAL STUDIES, AND THE LAW (Austin Sarat & Jonathan Simon eds., 2003) (noting that in several legal subfields, “culture” has replaced “social,” “public opinion,” and even “law”).

142. Human rights norms have, by design, an intuitive appeal. See LYNN HUNT, INVENTING HUMAN RIGHTS (2007) (tracing a brief history of the idea of universal human rights and noting that in order to succeed the concept needed to be natural).

143. For example, Sally Engle Merry has done important anthropological work on how global human rights norms become translated into local values in the context of violence against women. See SALLY ENGLE MERRY, HUMAN RIGHTS & GENDER VIOLENCE: TRANSLATING INTERNATIONAL LAW INTO LOCAL JUSTICE (2006).

144. Harold Hongju Koh, *The 1997 Frankel Lecture: Bringing International Law Home*, 35 HOUS. L. REV. 623, 626 (1998).

smoking, to start jogging, or to pay their taxes is not to coerce them, but to convince them to think of themselves as religious, non-smokers, joggers, or taxpayers.”<sup>145</sup> However, Koh does not turn to the literature on the micro-processes of norm promotion and adoption, a literature rife with prescriptions for would-be regime designers. While noting the success of campaigns to get people to “stop smoking, to start jogging, or to pay their taxes,” Koh does not cite to the communications literature evaluating the contexts in which those campaigns failed or succeeded.<sup>146</sup> But this literature can sharpen his theory. While he proposes a mechanism whereby norm entrepreneurs mobilize or build a transnational issue network that subsequently works with “interpretive communities” and law-declaring fora to bring about domestic political and subsequently legal change through norm internalization, recent norms literature suggests that interpretive communities could in fact endeavor to bring about behavior change without any norm-internalization.<sup>147</sup>

Few have heeded Koh’s call for what he calls the “second generation” of international law and international relations research that focuses on the micro-processes of social influence.<sup>148</sup> Evaluating the literature spawned by Koh’s theory, Rosa Brooks decries that “only a few legal scholars have addressed the question of how purposive governmental or nongovernmental norm-creation projects might actually work. This lacuna is even more noticeable in the domain of international and comparative law.”<sup>149</sup>

Engineering a system for human rights norm management is a tempting proposition.<sup>150</sup> Norms perform a social regulatory function without the en-

145. *Id.* at 628–29 (1998).

146. *Id.* For an example of how this literature can be usefully applied in legal scholarship, see Lessig, *The Regulation of Social Meaning*, *supra* note 24, at 1028, (citing Robert A. Kagan & Jerome H. Skolnick, *Banning Smoking: Compliance Without Enforcement*, in *SMOKING POLICY: LAW, POLITICS, AND CULTURE* 69, 82 (Robert L. Rabin & Stephen D. Sugarman eds., 1993)).

147. Koh, *supra* note 144, at 647 (calling transnational norm entrepreneurs “individuals who, in Ethan Nadelmann’s words, (1) ‘mobilize popular opinion and political support both within their host country and abroad’; (2) ‘stimulate and assist in the creation of like-minded organizations in other countries’; (3) ‘play a significant role in elevating their objective beyond its identification with the national interests of their government’; and (4) often direct their efforts ‘toward persuading foreign audiences, especially foreign elites, that a particular [normative] regime reflects a widely shared or even universal moral sense, rather than the particular moral code of one society’”) (citing Ethan A. Nadelmann, *Global Prohibition Regimes: The Evolution of Norms in International Society*, 44 *INT’L ORG.* 479, 482 (1990)). For support that behavior can align with a set of norms without any norm internalization, see BICCHIERI, *supra* note 20, at 3.

148. See Koh, *supra* note 144.

149. Rosa Brooks, *The New Imperialism: Violence, Norms, and the “Rule of Law,”* 101 *MICH. L. REV.* 2275, 2325 (2003).

150. See BICCHIERI, *supra* note 20, at 7 (“The idea that social norms may be cued, and hence manipulated, is attractive. It suggests that we may be able to induce pro-social behavior and maintain social order at low cost. Norms differ in different cultures, and what cues a Westerner into cooperation will probably differ from what cues a Mapuche Indian. In both cases, however, it may be possible to structure the environment in a way that produces desirable behavior.”) (citation omitted).

ergy or cost required by coercive enforcement mechanisms.<sup>151</sup> But because the scholarship has not yet outlined the social conditions where such an approach would work, norm management—and its behavioralist correlate, the management of social situations—is difficult to prescribe with any confidence. For example, work that suggests that norms have effect because they are “internalized” is contradicted by the work of social norms scholars that suggests that norms have an effect on behavior without any commitment from the individual whose behavior changes.<sup>152</sup> Behavioral insights are a necessary precursor to behavioral prescriptions, and the scholarship has not paid sufficient attention to those insights.

### B. *International Human Rights Practice*

Human rights practitioners, meanwhile, have pursued many innovative tactics for enforcing international human rights law, but their approaches tend to be preoccupied with rule-enforcement against individual violators,<sup>153</sup> rather than changing or constructing effective normative environments that would enable that rule enforcement. This is well displayed by the practice of shaming rights violators, a practice that is highly dependent on social context.

The classical account of human rights enforcement features actors—whether state or nonstate, and usually from the West but sometimes only tied to the West by funding or training—who investigate reports of violations of human rights laws. After documenting the abuses, they then press the perpetrators to change their behavior, threatening them with the shame of publicly exposing the documented harms, pressing for criminal or civil sanctions, or both.<sup>154</sup> The model has been enormously successful in some respects, perhaps most observably in changing the lexicon of legitimate international action.<sup>155</sup>

We might call this the “speak truth to power” model because that phrase, so-often used by rights practitioners, captures the implicit and

---

151. See Lessig, *The New Chicago School*, *supra* note 24, at 662 (“Norms constrain an individual’s behavior, but not through the centralized enforcement of the state. If they constrain, they constrain because of the enforcement of a community. Through this community, they regulate.”).

152. See BICCHIERI, *supra* note 20.

153. An international movement designed to “end impunity” has led to the creation of international criminal tribunals to punish bad actors who commit grave human rights abuses. For an account of one international criminal tribunal, see Craig Timberg, *Sierra Leone Special Court’s Narrow Focus: Well-Funded but Selective War Crimes Probe Draws Resentment of Impoverished Victims*, WASH. POST, Mar. 26, 2008, at A11.

154. See Emilie M. Hafner-Burton, *Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem*, 62 INT’L ORG. 689, 689–90 (2008) (noting that the tactic of naming and shaming is “common-place”). Of course this is not the only model for advocacy—advocates are creative and have designed many mechanisms for achieving their political goals. Nor is this model offered as a straw-man. This model of advocacy is a dominant one. It is still taught at law schools across the West and is practiced by the biggest and most powerful human rights organizations.

155. See generally Hafner-Burton, *supra* note 154 (noting the success human rights advocates have had in changing the language of international relations, but also noting the failure of that change to relate to a significant shift in state policies).

widely-held sentiment that truth *is* power. The model is premised on an Enlightenment-era conception of reason that assumes people, when confronted with certain information, will process that information through a cognitive function and return a rational, ergo human rights-sympathetic, response.<sup>156</sup> This approach is reflected in the language used to describe this work: human rights advocates are called “watch-dogs,” and their research trips are called “fact-finding” missions.<sup>157</sup> When those facts are reported, the tactic is called “shame-and-blame,” because the fact of a rights violation is assumed to evoke those two responses. As the language makes clear, the goal is to shine a light on whatever wrongs have occurred and then wait for the outrage. Kenneth Roth, Director of Human Rights Watch, describes the methodology this way:

The essence of that methodology . . . is not the ability to mobilize people in the streets, to engage in litigation, to press for broad national plans, or to provide technical assistance. Rather, the core of our methodology is our ability to investigate, expose, and shame.<sup>158</sup>

The core of the methodology is not to change social situations; rather, it is to shame human rights violators. But from where does this shame emerge?

The specialty of Roth and his colleagues is to scrupulously measure behavior against international norms and announce any deviations from those norms. The shame is presumed to come from the deviations. But, as Roth acknowledges, deviations from international norms do not always produce shame. He goes on to suggest that the shame-and-blame tactic works best where there is a clear rights violation, a clear violator, and a clear remedy.<sup>159</sup> Implicit in Roth’s argument is an acknowledgement of the power of social situations: in some contexts, practices that deviate from international human rights norms produce shame, and in others they do not.

When shaming tactics work—as in, for example, the Abu Ghraib prison abuse scandal, where lawyers and the public shamed the U.S. military for mistreatment of prisoners, eventually causing the closure of that prison—they likely do so because they harness the social situation around the target actor, which is possible because the harms documented strike a chord with the relevant audience.<sup>160</sup> The action that the audience is urged to demand is

---

156. See, e.g., KERRY KENNEDY CUOMO, *SPEAK TRUTH TO POWER* (2000) (profiling advocates who exposed human rights violations).

157. See HANS THOOLEN & BERTH VERSTAPPEN, *HUMAN RIGHTS MISSIONS: A STUDY OF THE FACT-FINDING PRACTICE OF NON-GOVERNMENTAL ORGANIZATIONS* (1986).

158. Kenneth Roth, *Defending Economic, Social and Cultural Rights: Practical Issues Faced by an International Human Rights Organization*, 26 *HUM. RTS. Q.* 63, 67 (2004).

159. See Roth, *supra* note 158.

160. For a discussion of the reaction to the abuse allegations, see Sewell Chan & Thomas E. Ricks, *Iraq Prison Supervisors Face Army Reprimand; Probe of Interrogations May Bring More Changes*, *WASH. POST*, May 4, 2004, at A1.

state enforcement of the international norm. But that model plainly depends on a social context that will enable the shame to occur.

The “truth” of the harms is not always so self-evident, and in those cases shaming does not work. Recall what happened when Human Rights Watch sought to apply its methodology to the troubling rise of civilian casualties during the conflict in Afghanistan in 2007.<sup>161</sup> In that particular case, one of the researchers decided to try something new, a deviation from standard practice, and a tactic that involved consciously squaring advocacy strategy with the local social reality. When the radio aired the family stories of loss due to suicide bombings, the Taliban’s social situation—that is, their position in Afghan society—had been activated, and the violation of international norms suddenly had been infused with local meaning. The Taliban, who seemed impervious to the shame of violating international human rights, suddenly risked shaming from Afghans. This is especially striking because previous appeals that treated the Taliban as a unitary actor—one that could be persuaded to care about human rights through the documentation of its rights abuses—had failed. The shame was socially contingent. Advocates know this all too well, and choose their campaigns and campaign language selectively depending on the social context.<sup>162</sup> What is missing is a formal account of how social situations function to regulate behavior in the human rights regime and how they can be harnessed and changed rather than merely navigated through.

Roth’s model acknowledges but does not account for the importance of the social factors that determine when shaming tactics work and when they do not, and he offers no systematic attempt to shape those factors to enable his shaming tactics. Instead, he concludes that shaming tactics work most effectively in situations of civil and political rights violations because in those situations the shaming tactics can be directed against a rights violator.<sup>163</sup> This causes him to further conclude that human rights groups engaged in the shame-and-blame tactic should focus on civil and political rights more than economic, social, and cultural ones.<sup>164</sup> A behavioral approach, on the other hand, would counsel against this hierarchy of human rights norms, and if anything it might privilege economic, social, and cultural rights.<sup>165</sup> Moreover, precisely because the model does not start with an

---

161. See *supra* Introduction.

162. See Negar Azimi, *Prisoners of Sex*, N.Y. TIMES, Dec. 3, 2006, § 6 (Magazine), at 63 (noting the strategy of Human Rights Watch to couch homosexual abuse in the language of torture rather than sexual equality).

163. See Roth, *supra* note 158, at 68 (noting that Human Rights Watch’s methodology relies on identifying a clear rights violation and a remedy, and that “[t]hese answers flow much less directly from the mere documentation of an ESC rights violation than they do in the civil and political rights realm”).

164. A related and significant limitation of this model—and this is not to deny its potency up until now—is its preoccupation with shaming violators, who are taken to be unitary actors, rather than social situations of rights violations. Such a focus necessarily limits the scope of action that human rights advocates can address.

165. See *infra* Part III.A.

explicit concern for social situations, it treats shaming as the only social mechanism for shaping behavior when in fact there are many social tools—disgust, for example—that operate in any given situation to influence individual behavior.<sup>166</sup>

### C. Modeling Compliance

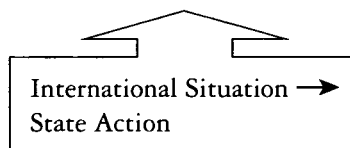
Some international human rights scholarship reflects the power of social situations, but only in the context of creating international political situations that socialize states into following the behavior of like-minded states. Meanwhile, some international human rights practitioners acknowledge the power of situations, but only insofar as the situation will activate a norm to get state enforcement of a particular rule. Neither scholarship nor practice is explicitly concerned with the power of local social situations to directly regulate the behavior that will eventually come to be the violation or promotion of human rights.

We can depict these approaches visually. If we think of international human rights laws as containing a norm and an enforcement mechanism, a simplistic model of international law might look like the following:

**International Norm + Enforcement Mechanism  
= International Law**

Debates about the robustness of international human rights law have generally turned on claims about the enforcement mechanisms in place at the international level, rather than debates about the strength of international norms. In the above model, for example, Goldsmith and Posner do not see a convincing enforcement mechanism and so they do not see robust international law.<sup>167</sup> Goodman and Jinks, on the other hand, see a potentially compelling enforcement mechanism in the form of state acculturation at the international level.<sup>168</sup> Their vision of enforcement can be depicted as the following:

**International Norm + Enforcement Mechanism  
= International Law**



166. See *infra* Part II.C.

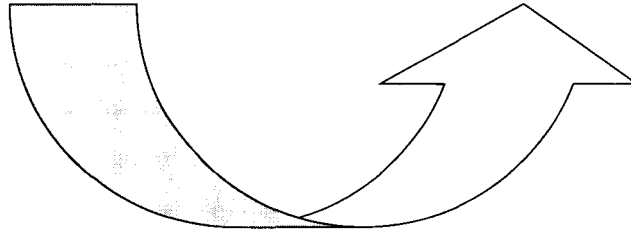
167. See GOLDSMITH & POSNER, *supra* note 118.

168. See Goodman & Jinks, *supra* note 129, at 646–56.

But as Koh notes, a problem with this acculturation model is that the international norm is not necessarily required for creating compliance.<sup>169</sup> There is a risk of decoupling where a state makes a commitment because of its international social situation without any real concern for or purchase of the related norms. The result is that the norms embodied by the international law may not “stick” and the domestic situation may not change.

Finally, consider Kenneth Roth’s account, which many human rights practitioners use.<sup>170</sup> It depends a great deal on the chance that under present conditions a violation of the relevant international norm will create an opportunity to shame the violator-state:

**International Norm + Enforcement Mechanism  
= International Law**



<i>International Norm Violation</i> + (Certain) Situations <hr style="width: 50%; margin: 0 auto;"/> = Shame
--------------------------------------------------------------------------------------------------------------------

The efficacy of the model depends on the situations that activate the relevant norms; the model does not actively seek to understand or cultivate that norm-situation. Nor do practitioners actively seek out alternative models that are only imaginable with social situations in mind.<sup>171</sup>

As these simple models show, the dominant accounts of international human rights regimes implicitly acknowledge the importance of social situations, but they do not take them into account as explicit regulatory tools; none of the models treat local social situations as the crucial units of regulatory analysis.

---

169. See Koh, *Internalization Through Socialization*, *supra* note 134, at 981.

170. See Roth, *supra* note 158.

171. See *infra* Part III.C.2.



### III. IMPLICATIONS FOR HUMAN RIGHTS

Anyone in a modern market economy knows that actors can and do consciously construct situations to influence behavior. Nearly every multinational grocery store in the world reflects a set of design features meant to regulate customer behavior.<sup>172</sup> These design decisions are based on a set of insights about consumer behavior: some shoppers are naturally impulsive;<sup>173</sup> active music and lighting reduce cognitive function and increase reliance on affect, hence boosting sales of affect-laden products;<sup>174</sup> and while choosing from a large assortment of goods can lead a consumer to make a more “virtuous” or utilitarian decision than a smaller selection would, these findings reverse in the presence of situational factors that might justify indulgence.<sup>175</sup>

Just as fundamental marketing principles eventually changed the design of the grocery store, so too may behavioral insights offer the prospect to rethink the design of international human rights regimes. Constructing a comprehensive regime is beyond the scope of this Article. Instead, this Part examines a few key human rights regime design choices in a behavioralist light. The specific questions broached are examples of the sort of re-imagining that is made possible by behavioral insights. These are not fully tested applications but rather provocations and calls for future research—signposts in a research agenda that will take many years to fill out.

Moreover, these are not suggestions for a radical rethinking of the human rights regime: for instance, they do not explore non-legal measures. Instead, this Part explores how certain features of today’s international human rights regime might look different in order to reflect a concern not only with unitary actors—whether states or individuals—but also with the relevant social situations.

#### A. *The Content of Human Rights Laws*

This section discusses the implications of behavioral insights about the power of social influence for three aspects of human rights law: the social conditions that make it likely that state actors will comply with their obligations; debates about whether human rights regimes should focus on grave abuses of human rights or smaller breaches that may contribute to social

---

172. See JAMES M. MAYO, *THE AMERICAN GROCERY STORE: THE BUSINESS EVOLUTION OF AN ARCHITECTURAL SPACE* 242 (1993). See also Roger March & Arch G. Woodside, *Advancing Theory on Consumer Plans, Actions, and How Marketing Information Affects Both*, in *APPLYING SOCIAL COGNITION TO CONSUMER-FOCUSED STRATEGY* 155, 156 (Frank R. Kardes et al. eds., 2005).

173. See Sharon E. Beatty & M. Elizabeth Ferrell, *Impulse Buying: Modeling its Precursors*, 74 *J. RETAILING* 169, 184–85 (1998).

174. See Baba Shiv & Alexander Fedorikhin, *Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making*, 26 *J. CONSUMER RES.* 278, 289 (1999) (noting that impulse buying will go up in environments where deliberative cognition is limited).

175. See Aner Sela et al., *Variety, Vice, and Virtue: How Assortment Size Influences Option Choice*, 35 *J. CONSUMER RES.* 941, 942 (2009).

conditions that will make grave abuses more likely; and finally, debates about the relevance and viability of economic, social, and cultural rights.

### 1. *Creating the Conditions of Compliance*

Human rights scholars are deeply concerned with how to enforce human rights treaties against recalcitrant regimes.<sup>176</sup> They occasionally discuss embedding diplomats—and, by extension, states—in social situations that will increase compliance. However, the scholarship rarely turns to discuss devices designed to harness domestic social situations, rather than political constituencies, to increase the resonance of international human rights norms where they matter most.

Jacob Gersen and Eric Posner, expressing a common sentiment, note that “[w]riting down our moral views on a piece of paper should not make them any stronger.”<sup>177</sup> But this is contradicted by recent scholarship. As norms scholar Christina Bicchieri explains, after reviewing the relevant empirical literature, “The very act of promising, ‘cheap talk’ of no consequence, might be enough to induce many of us to behave contrary to narrow self-interest. A social norm has been activated, and, under the right circumstances, we are prepared to follow it.”<sup>178</sup> Such a finding could be used to justify building into human rights laws a requirement that every public official responsible for enforcing an international human rights obligation make public and written commitments with the express aim of embedding them in social structures that will bind them to those commitments.

This is a step beyond the conventional account of how a leader makes a commitment to a set of norms—for instance the Helsinki Accords<sup>179</sup>—and how those commitments grow legs. Rather, drawing from the commitment/consistency bias, international treaties could include a requirement that all state actors literally write down their commitment to the international obligation and also publicly proclaim to their family and community their support of such a commitment. These “cheap talk” commitments are low-cost and potentially high-gain, especially for the realist who is unaware of the power that their social situation holds over their behavior. An even more ambitious—and maybe less hokey—commitment strategy would be to require all signatories to launch a media campaign heavily publicizing their international commitment domestically. One way to achieve this would be for the agreement to authorize an outside body to publicize the international commitment domestically, thereby circumventing the question of whether a

---

176. See *supra* Part II.A.

177. Jacob E. Gersen & Eric A. Posner, *Soft Law: Lessons from Congressional Practice*, 61 STAN. L. REV. 573, 592 (2008).

178. BICCHIERI, *supra* note 20, at 175.

179. See Ryan Goodman and Derek Jinks, *Incomplete Internalization and Compliance with Human Rights Law*, 19 EUR. J. INT'L L. 725, 739 (using the Helsinki Accords as an example of how political leaders are affected by their public commitments).

recalcitrant regime would follow through with its duty to make its commitment public.

There are obvious concerns about whether such designs will work. But despite a great deal of consternation in the international legal literature over the effect of what seem like empty commitments,<sup>180</sup> several legal regimes already reflect an intuitive understanding of the commitment/consistency bias's power. Contract theory accounts for the finding that promises take on new meaning once made, despite initial intentions.<sup>181</sup> The requirement that police read suspects their *Miranda* rights is designed to deliver a set of information, but also and perhaps more importantly to place police officers in a situation where their behavior is governed by the words they speak.<sup>182</sup> During the American invasion of Iraq, detainees who had been cleared for release from detention were frequently asked to take an oath in front of their family that they renounced violence.<sup>183</sup> And public officials in the United States are asked to pledge an oath to uphold the Constitution in their work—a duty many acknowledge has subtle but profound effects on their behavior.<sup>184</sup>

But even if the design will work, are states likely to agree to such a commitment? The paradox of international law is that recalcitrant states are unlikely to agree to strong enforcement mechanisms that they fear will bind them. Behavioral insights help us here too. First, states already make commitments simply by signing human rights agreements. And they follow up those commitments with re-commitments such as the 1993 World Conference on Human Rights, in which 171 states adopted the Vienna Declaration and Programme of Action.<sup>185</sup> That program included a recommendation that states draw up a “national plan identifying steps whereby that state would improve the promotion and protection of human rights.”<sup>186</sup>

Such a plan is a good beginning, but a behavioral approach might improve it by simply requiring that states not only draw up action plans but also take the first step in their action plan. As noted above, “channel factors” are influential situational forces.<sup>187</sup> For example, interview subjects who were informed of the benefits of a tetanus inoculation shot were highly

180. See GOLDSMITH & POSNER, *supra* note 118, at 175–79 (describing how much of international legal rhetoric is “cheap talk”).

181. See Robert A. Prentice, *Law & Gratuitous Promises*, 2007 U. ILL. L. REV. 881, 881 (2007) (“Behavioral law and economics tends to show that the same limits on rational human decision making that apply to gratuitous promises also apply to bargained-for commercial exchanges.”).

182. See Charles D. Weisselberg, *Mourning Miranda*, 96 CAL. L. REV. 1519, 1527–28 (2008).

183. See Alexandra Zavis, *The Conflict in Iraq: A Different Approach*, L.A. TIMES, Nov. 16, 2007, at A12..

184. See JACK GOLDSMITH, *THE TERROR PRESIDENCY* 28 (2007) (“The oath of office and a powerful professional concern to ‘do the right thing’ help OLC lawyers to resist pressures for certain outcomes when they believe the law requires otherwise.”).

185. See The Secretary-General, *Report of the World Conference on Human Rights: Report of the Secretary General*, U.N. Doc. A/CONF.157/24 (Oct. 13, 1993).

186. World Conference on Human Rights, June 14–25, 1993, *Vienna Declaration and Programme of Action*, ¶ 71, U.N. Doc. A/CONF.157/23 (July 12, 1993).

187. See *supra* text accompanying note 49.

unlikely to actually get those shots; only three percent eventually got inoculated. But when students “received the same communication but were then also given a map of the campus with the infirmary circled, and urged to decide on a particular time and route to get them there,” the rate increased to twenty-eight percent.<sup>188</sup> Seeing the mapped route (channel factor) and setting a specific time (first-step commitment) significantly affected individual behavior. The information alone was not enough to get subjects to act, but a small amount of context significantly enhanced the information’s guiding effect on the subjects’ behavior. Human rights action plans could seek to harness the same channel factors—both to affect state behavior and street level bureaucrat behavior.

The risk to such an approach is that the first step requirement may deter some states from making a commitment at all if they fear it will lead them down a commitment-laden path. But as previously noted, behavioral insights suggest that public officials are likely to underestimate the power of these strengthened socially-contextual commitments. The fundamental attribution error—which holds that humans tend to overemphasize the role of disposition in explaining others’ actions and underemphasize the role of the situation—suggests that the circumstances that activate certain norms are in fact much more powerful than we think.<sup>189</sup> That is to say: we underestimate the power of situational effects, and states may do so as well. Public officials making this error could plausibly sign onto a commitment underestimating its potency. It is a much smaller requirement to ask a state to publicize a set of rights than to enforce them with the full panoply of law-enforcement machinery. This also allows states to tailor their responses to their capacity, rather than having to accept a commitment that they cannot uphold.<sup>190</sup> This indirect approach may appeal to the state and may be more likely to be successful if the state underestimates the power of the social situation to regulate its behavior. This is a low-cost addition to a human rights agreement, and the potential payoff is significant.

## 2. *Minor Abuses or Grave Breaches?*

There is a debate among human rights and humanitarian lawyers about whether to focus attention on grave or minor abuses.<sup>191</sup> Behavioral insights

---

188. Marianne Bertrand et al., *supra* note 49, at 6.

189. See Lee D. Ross, *supra* note 28.

190. See ABRAM CHAYES & ANTONIA CHAYES, *THE NEW SOVEREIGNTY: COMPLIANCE WITH INTERNATIONAL REGULATORY AGREEMENTS* 10 (1995) (suggesting that some of the most significant limitations on state compliance with international law are “limitations on the capacity of parties to carry out their undertakings”).

191. For a helpful discussion of this debate in the context of international criminal law, see Luc Côté, *Reflections on the Exercise of Prosecutorial Discretion in International Criminal Law*, 3 J. INT’L CRIM. JUST. 162, 169 (describing the level of prosecutorial discretion in international criminal law with regard to indictments in terms of both numbers and rank of the indicted, and noting the criticism which “was directed at the Prosecutor for the indicting lower perpetrators, considered ‘small fries,’ particularly in the first cases of *Tadić* at the ICTY and *Akayesu* at the ICTR.”).

cut both ways and may suggest a combined approach. On the one hand, the fact that a particular social situation can lead to rights abuses may counsel for criminal sanctions against the highest authorities in charge of designing the institutions that facilitate such situations. At the same time, the same set of research insights suggests that those situations can be avoided through careful monitoring and oversight of minute details—a finding that calls for close monitoring and sanctions over the smallest infractions in order to create a general tone of law-abidingness.

In the now well-known Stanford Prison experiment, psychologist Philip Zimbardo showed how otherwise-peaceful students could begin to act in terribly abusive ways—so abusive, in fact, that his experiment was aborted.<sup>192</sup> As he describes it, many of his ‘prison guards’ engaged in degrading behavior:

My guards repeatedly stripped their prisoners naked, hooded them, chained them, denied them food or bedding privileges, put them into solitary confinement, and made them clean toilet bowls with their bare hands . . . . Over time, these amusements took a sexual turn, such as having the prisoners simulate sodomy on each other . . . . Human behavior is much more under the control of situational forces than most of us recognize or want to acknowledge.<sup>193</sup>

Perhaps most striking about Zimbardo’s experiment was that his guards were well-educated and empathetic Stanford students from the same community as his prisoner subjects; they were not in a war zone; and they were in a simulated prison, not a real one. If we can assume that his subjects were not particularly corrupt by nature, then the experiment—and other examples of gross abuse—raises the challenge not of convincing people to be good but rather of “trying to arrange social institutions so that human beings are not placed in situations in which they will act badly.”<sup>194</sup> That is, the behaviorist approach may call for a focus not only on the criminality of action at hand—which can be left for a separate determination—but also for a focus on the following pragmatic questions: how did such a situation come about, and how can it be prevented in the future? This evaluation of the *ex ante* situation is different from the conventional human rights approach of identifying violators and calling for their punishment.

Again behavioral insights help us to answer these questions. Psychologist Craig W. Haney, who served as a lead researcher on the prison experiment

---

192. See Craig Haney et al., *Interpersonal Dynamics in a Simulated Prison*, 1 INT’L J. CRIMINOLOGY & PENOLOGY 69, 88 (1973).

193. Philip G. Zimbardo, *Power Turns Good Soldiers Into ‘Bad Apples’*, BOSTON GLOBE, May 9, 2004, available at [http://www.boston.com/news/globe/editorial\\_opinion/oped/articles/2004/05/09/power\\_turns\\_good\\_soldiers\\_into\\_bad\\_apples/](http://www.boston.com/news/globe/editorial_opinion/oped/articles/2004/05/09/power_turns_good_soldiers_into_bad_apples/).

194. KWAME ANTHONY APPIAH, EXPERIMENTS IN ETHICS 71 (2008).

described above and now works on the application of psychological insights to prison policy, has noted that all that is required to prevent detainee abuse is “regular training and discipline, along with outside monitoring.”<sup>195</sup> This idea can be clearly seen in practice. For example, interviews with soldiers and the leadership of Camp Bucca suggest that the camp underwent a process of deliberate “professionalization”—that is, strict discipline among the soldiers that included attention to seemingly inconsequential details such as the cleanliness of the gravel roads.<sup>196</sup> According to the commanding general, Major General Doug Stone, the aim was to create a social situation in which the guards’ behavior was almost pre-determined to be rule-abiding.<sup>197</sup> For example, the guards regularly held competitions over which detainee compound was the cleanest—a competition that harnessed social pressure in a race toward norm compliance.<sup>198</sup> The monitoring of these micro rule violations, the theory goes, contributed to a cultural shift that led to self-enforcement and prevention of grave abuses.

A recent report by the International Committee of the Red Cross (“ICRC”) supports this idea.<sup>199</sup> The report, which included one historical review and three surveys of combatants, civilians, and ICRC delegates, found that combatants exhibit high levels of “group conformity” and obedience to authority.<sup>200</sup> The report found that “[w]hile violations of IHL may sometimes stem from orders given by [superiors], they seem more frequently to be connected with a lack of any specific orders not to violate the law or an implicit authorization to behave in a reprehensible manner.”<sup>201</sup> The report concludes that punishing minor abuses is just as crucial, if not more so, than grave abuses.<sup>202</sup>

The conclusion for human rights regimes—in civilian as well as military situations—should be to pay more careful attention to minor abuses and situational factors that might contribute to a bandwagon of abuse. But due

195. John Schwartz, *Simulated Prison in '71 Showed a Fine Line Between 'Normal' and 'Monster'*, N.Y. TIMES, May 6, 2004, at A20.

196. Interview with Douglas Stone, Major General, United States Marine Forces Reserve, in Camp Bucca, Iraq (Jan. 8, 2008) (on file with the Harvard Law Library).

197. *Id.*

198. *Id.*

199. Daniel Muñoz-Rojas & Jean-Jacques Frésard, *The Roots of Behaviour in War: Understanding and Preventing IHL Violations*, 86 INT'L REV. RED CROSS 189 (2004).

200. *Id.* at 193–94.

201. *Id.* at 194.

202. *See id.* at 204.

Any failure to obey an order must be sanctioned. Sanctions, which are central to determining a combatant's behaviour, can take different forms (e.g. disciplinary, penal or social). Disciplinary or penal sanctions should be promoted both for the purpose of setting an example and for prevention. It is essential that the authorities should take action, even for offences which are less serious than a war crime, so as to ensure the discipline of their troops and avoid entering a spiral of violence in which violations may become not only more and more serious but also more and more acceptable in the eyes of those who commit them.

to resource constraints as well as the institutional and professional prerogatives of the current human rights regime, insufficient attention is paid to those situational factors. Take, for example, an international criminal tribunal established long after the seeds of conflict have blossomed into a civil war. By the time of the tribunal, it may be too late to address structural causes of human rights abuses. From the *ex ante* position, a behavioral approach would draw more heavily on careful monitoring of minor rights abuses before they contribute to what the ICRC study calls “the spiral of violence.”<sup>203</sup>

However, if we take criminal trials to be useful mechanisms for deterring future action, such a finding may also call for criminalizing the creation of social situations of harm. Thus, for example, even though Julius Streicher is not hanged for minor abuses, he is convicted for his misuse of authority while in a position to greatly influence ground-level social situations that contribute to great harms.<sup>204</sup> In this way, while behavioral insights counsel for attention to minor abuses that may snowball into a situation of grave abuses, those same insights counsel for attention to grave abuses by those in a position of authority to create situations that contribute to the spiral of violence.

### 3. *Economic, Social and Cultural Rights*

The human rights community has long been divided over whether to pursue and promote economic, social, and cultural rights (“ESCRs”) with the same vigor as civil and political rights (“CPRs”).<sup>205</sup> This is the result of many concerns including distributional fears, cultural sensitivities, and a question about the ability to measure and enforce ESCRs. But behavioralism suggests that this debate is incomplete and, more importantly, offers a new way of discussing the terms of the debate.

As noted above, leading human rights lawyer Ken Roth has argued passionately that ESCR work does not fit well with the expose-and-shame model of advocacy because ESCR cases contain more uncertainty than CPR cases.<sup>206</sup> But by taking a behavioral approach to, say, the right to health in Ghana, we immediately see that there is no easily identifiable unitary violator. Instead, there is a rights-negative situation and the solution to the “rights violation” lies in the behavior of ordinary Ghanaians. The solution, at least with regard to this behavioral issue, is a purposeful shift in the social meaning of handwashing. Once ESCRs are seen as guidelines for what sorts of social situations to engender, rather than guidelines for who should be punished and why, they may become more manageable—and as a result, more readily a part of a productive human rights practice.

---

203. See Muñoz-Rojas & Frésard, *supra* note 199, at 195.

204. See IMT, *supra* note 1.

205. See INTERNATIONAL HUMAN RIGHTS IN CONTEXT 263 (Philip Alston et al. eds., 3d ed. 2007).

206. Roth, *supra* note 158, at 68.

Behavioral insights also destabilize the rigid conceptual walls between civil and social rights. A behavioralist approach to human rights may even counsel for privileging certain ESCRs over CPRs, where those ESCRs—a right to education, certain cultural practices, and economic opportunity—enable the social situation to function as a medium for other human rights. Without those ESCRs, it may be impossible to create rights-positive situations, for they depend on privileging certain pieces of information over others or a certain approach to that information over others. Efforts to establish freedom of speech, for instance, may focus a disproportionate amount of time on laws that limit speech rather than on the economic, educational, and technological barriers to true freedom of speech. The mere absence of state censorship does not automatically create agency of speech. Indeed, access to the right networks and markets is important for enabling situations that engender speech rights even in the face of government censorship. This is well illustrated in the case of Facebook in Egypt: new technologies can, in some cases, allow transnational advocacy networks to communicate while circumventing state censorship.<sup>207</sup>

Behavioral insights do not necessarily suggest that norm hierarchies be reversed, but they do shift scholarly and practical focus—perhaps toward innovative ways to create norm-situations that have little to do with distinctions between first and second generation rights.

### *B. Adjudicating Human Rights*

The fundamental insights of behavioralism have specific implications for design choices by the institutions established to interpret human rights law, and also for the content of their interpretations of human rights law. This brief application addresses two questions: first, what are the implications of behavioral insights for debates about the mechanisms for interpreting human rights law in the context of transitional justice? Second, can behavioral insights be deployed by an adjudicatory body to better define human rights crimes?

#### *1. Transitional Justice Mechanisms*

Behavioral insights by no means settle the debate about the form that transitional justice should take—a debate that is often framed as one of “truth versus justice.”<sup>208</sup> Rather, they suggest positive and negative aspects

---

207. See, e.g., Hiawatha Bray, *Beating Censorship on the Internet*, BOSTON GLOBE, Feb. 20, 2006, at A10. (describing the TOR service that masks a computer's identity online). But see JACK L. GOLDSMITH & TIM WU, WHO CONTROLS THE INTERNET: ILLUSIONS OF A BORDERLESS WORLD (2006) (debunking utopian visions of an unfettered, borderless internet).

208. See generally PRISCILLA B. HAYNER, UNSPEAKABLE TRUTHS: CONFRONTING STATE TERROR AND ATROCITY 10–23 (2001) (describing the differences between a truth commission and a criminal tribunal).



of the dominant mechanisms of transitional justice, and may prove useful in both contexts.

A behavioral approach to transitional justice supports the role that truth commissions play in identifying the social foundations of a particular conflict. Likewise, the recommendations that truth commissions make for how to remedy the harms of social destruction are often explicitly behavioral. But behavioralism may also counsel for criminal trials if and where they punish social regulators, as in the Streicher trial, while it may counsel against criminal trials that punish the ground level actors who commit abuses when those abuses were highly predetermined by situational factors. Behavioralism may thus offer another layer of analysis in the debates about what sort of transitional justice mechanism to institute. Once that mechanism is in place, behavioral insights may further be useful for interpreting and applying human rights law to serve the desired ends.

A behavioral approach to transitional justice would ask whether any particular institution is sufficiently focused on the relationship between society and behavior and, just as importantly, how the creation of that institution shapes the situation on the ground. The fundamental attribution error cautions against the natural tendency to place blame for situational disasters on a small number of so-called bad actors—especially where such punishment diverts resources and attention from the ultimate goal of preventing such situations in the future. Similarly, a behavioral approach would ensure that any punishments did not crowd out parallel projects related to justice. Such an approach might be extremely useful in determining, for example, the choice between a truth commission, a criminal tribunal, or some other device.

After the bloody civil war in Sierra Leone, states and non-state actors from around the world contributed time, expertise, and money to establish the Special Court for Sierra Leone. That court is housed in a massive courthouse in Freetown, where, to date, five men have been convicted at a cost of \$150 million.<sup>209</sup> Meanwhile, the situation that led to the civil war in the first place—the shadow of colonialism, poor education levels, poor health, a culture of corruption, and vast economic inequalities—remains largely unchanged.<sup>210</sup> The court's attribution of the crimes of the civil war to these five men has not salved the wounds, let alone the causes, of the civil war.<sup>211</sup> A behavioralist approach to transitional justice might instead focus not only on just the direct causal links between certain actors and destructive actions but also on where, how, and why the social fabric enabled a killing field. The

---

209. Timberg, *supra* note 153. The court was also designed as an outreach and training mechanism, but it has been disappointing in that capacity. *See id.*

210. *See* Sahr Kpundeh, *Corruption and Political Insurgency in Sierra Leone*, in *BETWEEN DEMOCRACY AND TERROR: THE SIERRA LEONE CIVIL WAR* 90, 91 (Ibrahim Abdullah ed., 2004) (describing the country's colonial past and describing poverty and corruption as the "proximate causes" of the rebel war).

211. *See id.*

approach of the Special Court and a behavioralist approach may not be mutually exclusive; but the behavioralist approach allows us to see that situational conditions are not sufficiently addressed by current criminal tribunals alone.

The implications of behavioralism for transitional justice do not obviously counsel for holding criminal trials over establishing truth commissions. However, work on social networks may suggest that the approach of international criminal tribunals targeting a few key individuals is in fact the best way to disrupt the entire regime, even if such an approach does little in the way of addressing larger social dynamics. As Claire Dickerson notes, in a project that examines corporations as complex social networks:

A few social actors have sought to exert influence over corporations by a method that presupposes a very different conception of the corporation. These actors have targeted particularly well-connected individuals within the corporation, persons that the transaction-cost and social-responsibility perspectives fail to identify. By targeting these persons, the social actors have affected the entire enterprise, just as surely as an event that affected the electrical grid near Cleveland, Ohio for less than a second pulled the plug on fifty million people from Detroit to Toronto to New York City.<sup>212</sup>

We might imagine criminal tribunals making indictments based on a similar analysis. However, those indictments will vary significantly depending on whether the court's motivation is to respond to political or moral charges against certain vilified actors rather than, say, carefully-chosen nodes in a particular social network.

It is possible that behavioral insights will shift not only debates about truth versus justice but also debates about the form and shape of institutions within each mechanism. In the case of criminal trials, the shift may be in the theory of punishment. As social psychological evidence continues to undermine the notion of a "free will"—the basis for retributive punishment—some have predicted a move away from retributive and toward consequentialist theories of justice.<sup>213</sup> In the context of mass atrocity, classical psychology studies and more recent legal scholarship alike suggest that individual criminal punishments are an ill-suited response to the social dynamics of

---

212. Claire Moore Dickerson, *Corporations as Cities: Targeting the Nodes in Overlapping Networks*, 29 J. CORP. L. 533, 535 (2004).

213. See, e.g., Joshua Greene and Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 PHIL. TRANS.: BIOLOGICAL SCI. 1775 (2004) (arguing that advances in neuroscience will change understandings of free will and causation such that retributive forms of justice will seem arbitrary).

group crimes because they are premised on different understandings of social action.<sup>214</sup>

Ultimately, the most promising aspect of the behavioral approach is that it would transcend the false dichotomy between truth and justice and shift attention to more important distinctions. Regardless of whether the transitional justice mechanism in place is one focused on criminal punishment or collective truth-telling, a situation-based approach would counsel for careful attention to the context in which the tribunal operates and the ways in which the tribunal can address social conditions in transitional societies.

For many tribunals, this may mean that education and outreach are not just add-ons, but crucial pre-requisites for justice. Human Rights Watch, for example, found that the International Criminal Court's ("ICC") issuance of an indictment in the Ituri province of the Democratic Republic of Congo had a dual effect on Ituri residents.<sup>215</sup> On the one hand, the report found that in the wake of the Lubanga indictment—which included charges of child conscription—some citizens learned that recruitment and conscription of child soldiers is a crime.<sup>216</sup> But, the same study revealed that many residents in Ituri "expressed the opinion that the ICC's charges against Lubanga of enlisting, forcibly recruiting, and using children to actively participate in hostilities [were] not 'serious.'"<sup>217</sup> This suggests that while ICC indictments offer the opportunity for local education regarding human rights, if that education is ineffective the trial may not have any local resonance. From a behavioral perspective, holding a trial that leads Ituri residents to see human rights obligations as not serious is counter-productive. But this finding is not necessarily an indictment of the court—rather, it has implications for the way the court functions. Education and outreach, as the Human Rights Watch report notes, are crucial to the court's success.

In the end, appropriate questions about transitional justice may have much less to do with whether the tribunal is punitive or not, or domestic or international, than with how the tribunal makes itself relevant to existing communities, social networks, and social meanings. That is, the details of the design of such a tribunal may draw on behavioral insights, including: whether to consider collective sanctions; whether to maximize public opportunities to be heard; and whether to consider creative forms of punishment besides imprisonment, such as shaming or community service.

---

214. See DRUMBL, *supra* note 124, at 29–30 (discussing the work of legal scholars Laurel Fletcher and Harvey Weinstein, as well as psychologists Gustav Le Bon and Sigmund Freud).

215. HUMAN RIGHTS WATCH, COURTING HISTORY (2008), available at <http://www.hrw.org/en/reports/2008/07/10/courting-history?print>.

216. *Id.* at 68.

217. *Id.* at 64.

## 2. Developing Legal Doctrine

In 2003, the International Criminal Tribunal for Rwanda (“ICTR”) found three men—two radio broadcasters and one newspaper publisher, none of whom were found to have committed physically violent acts—guilty for incitement to genocide and crimes against humanity.<sup>218</sup> The broadcasters had urged listeners to slaughter Tutsis. Many listeners did: in the Rwandan genocide of 1994, Hutu civilians destroyed nearly one million Tutsi and Tutsi-sympathizers.<sup>219</sup> *Prosecutor v. Nahimana*, the so-called “media trial,” was an attempt to understand, for the purposes of international regulation, the social mechanisms underlying this media-coordinated genocide.<sup>220</sup>

In its opinion, the ICTR defined the crime of incitement to genocide through metaphor: even if the propagandists were not the trigger that set off the violence in the spring of 1994, they “were the bullets and the gun. The trigger had such a deadly impact because the gun was loaded.”<sup>221</sup> The genocide was a function of social dynamics, and those dynamics were a function of intentional media manipulation; that was enough, in the court’s eyes, to justify harsh criminal sanctions.<sup>222</sup> The court—weakly, according to some critics<sup>223</sup>—defends the lack of direct causation by referring to the defendant’s effects on the genocidal atmosphere:

The evidence does not establish a specific link between the publication [of *Kangura*, the Hutu-controlled newspaper whose publisher was on trial] and subsequent events, and yet such a link was clearly perceived by many witnesses such as Witness AHI, Witness ABE and Nsanzuwera, suggesting that *Kangura* greatly contributed to the climate leading to these events, if not causing them directly.<sup>224</sup>

The court does not clarify the standard of “contribut[ing] to the climate” of genocide, nor does it delineate how future courts might determine direct causation.<sup>225</sup> These propaganda crimes had not been revisited by an interna-

218. *Prosecutor v. Nahimana*, Case No. ICTR-99-52-T, Judgment and Sentence (Dec. 3, 2003).

219. *See supra* Part I.B.

220. *See Nahimana*, Case No. ICTR-99-52-T.

221. *Nahimana*, Case No. ICTR-99-52-T, at ¶ 953.

222. Two of the three men were originally sentenced to life in prison, see *Nahimana*, Case No. ICTR-99-52-T, at ¶¶ 1105–08 but on appeal all three sentences were reduced to between 30 and 35 years imprisonment, see *Nahimana v. Prosecutor*, Case No. ICTR 99-52-A, Judgment, ¶ 744 (Nov. 28, 2007).

223. See Joshua Wallenstein, *Punishing Words: An Analysis of the Necessity of the Element of Causation in Prosecutions for Incitement to Genocide*, 54 STAN. L. REV. 351 (2001).

224. *Prosecutor v. Nahimana*, Case No. ICTR-99-52-T, Judgment and Sentence ¶ 242 (Dec. 3, 2003).

225. *Id.*

tional tribunal since the Streicher trial at Nuremberg which, as the ICTR noted in procedural history, also had not required direct causation.<sup>226</sup>

Nowhere does the court apply any of the aforementioned social science insights to explain the media moguls' roles in the genocide. There is no mention of social proof or social networking. Instead, the court relies on hundreds of pages of testimony about the genocide from field researchers and concludes that the "climate" was poisonous.<sup>227</sup> The court's methodology does not lead to a robust rule of law and perhaps only works in the particular case because the roles of media were so prominent that the prosecutions were over-determined.

Instead, the court could have used behavioral insights to show that Hutu propagandists intentionally harnessed the social fabric to create a situation conducive to genocide. More specifically, the Hutu regime intentionally harnessed social proof to convince their listeners that they should join other Hutus in the killing; they intentionally manipulated the social meaning of killing a Tutsi (who had been labeled "cockroach," and an act that had been labeled "work"); and they intentionally constructed and manipulated social networks in order to control the flow of information to both Hutus and Tutsis.<sup>228</sup> Rather than relying on atmospherics to prove "incitement to genocide," the court could have used social science insights to make a more specific evidentiary requirement in order to prove the crime of incitement to genocide.<sup>229</sup>

Article III(c) of the Genocide Convention declares that "[d]irect and public incitement to commit genocide" is a crime.<sup>230</sup> ICTR case law defines incitement as communication recommending genocidal acts. As the court noted in the *Akayesu* trial:

[D]irect and public incitement must be defined for the purposes of interpreting Article 2(3)(c) [of the Statute of the Tribunal], as directly provoking the perpetrator(s) to commit genocide, whether through speeches, shouting or threats uttered in public places or at public gatherings, or through the sale or dissemination, offer for sale or display of written material or printed matter in public places or at public gatherings, or through the public display of placards or posters, or through any other means of audiovisual communication.<sup>231</sup>

226. See *id.* ¶ 981.

227. *Id.* ¶ 1073.

228. See *supra* Part I.B.

229. Prosecutor v. Nahimana, Case No. ICTR-99-52-T, Judgment and Sentence ¶ 953 (Dec. 3, 2003).

230. Convention on the Prevention and Punishment of the Crime of Genocide art. III(c), Dec. 9, 1948, 102 Stat. 3045, 78 U.N.T.S. 277.

231. Prosecutor v. Akayesu, Case No. ICTR 96-4-T, Judgment, ¶ 559 (Sept. 2, 1998).

The focus of such a definition is communication acts. The proof of such incitement is, as the appeals chamber in *Nahimana* found, in the particular communications meant to publicly incite genocide.<sup>232</sup> To make the determination that the defendants had committed the crime of incitement, the court reviewed particular pieces of speech and determined whether or not they contributed to the poisonous climate.

Instead of requiring hundreds of pages of evidence about the poisonous climate, the court could accept—or even require—as evidence of incitement certain findings about: the purposeful manipulation of social network structures; the purposeful harnessing of social proof; and the purposeful altering of the social meaning of an act of genocide thereby making it more likely.<sup>233</sup> Requiring that prosecutors prove at least one, or even all, of these elements would result in a more narrowly tailored prohibition of incitement to genocide, and might assuage some of the many concerns about that law being used to infringe protected speech.<sup>234</sup>

Secondly—and more ambitiously—the court could have used the social science evidence to find incitement in acts other than mere communication. That is, the court could find that a person is guilty of incitement to genocide when that person intentionally endeavors to create social conditions that make genocide possible by deliberately manipulating network structures, social proof, or social meaning to make genocide more likely. Some acts constituting incitement to genocide could be committed without ever uttering a word. Whereas traditional incitement looks at speech acts alone, a behavioral approach to incitement would examine whether social structures were manipulated in a way that makes genocide more likely to occur, regardless of whether someone made a direct and public statement that makes genocide more likely.

This idea, that behavioral insights may enhance legal doctrine, is not only applicable to incitement to genocide. Allison Danner and Jenny Martinez recently catalogued the jurisprudential roots of joint criminal enterprise liability, but they did not incorporate any psychological principles to explain the behavioral power or logic of the legal doctrine.<sup>235</sup> Our understanding of how accomplices are brought into criminal organizations may determine whether we think the law should require specific intent on the part of accomplices. If proof of the crime turns on, as Danner and Martinez suggest, not only the extent to which a defendant “made a substantial contribution to the fulfillment of the common objective of the enterprise”<sup>236</sup> but also the

232. *Nahimana v. Prosecutor*, Case No. ICTR 99-52-A, Judgment, ¶ 744 (Nov. 28, 2007).

233. This is not offered here as a perfectly defensible rule of law, but rather an example of a rule of law that is more specific than the one used in the court case thanks to its use of social science insights about situationism.

234. See Wallenstein, *supra* note 223.

235. See Allison Marston Danner & Jenny S. Martinez, *Guilty Associations: Joint Criminal Enterprise, Command Responsibility, and the Development of International Criminal Law*, 93 CAL. L. REV. 75 (2005).

236. *Id.* at 79.

defendant's intent, then the psychological evidence about the effect of group settings on individual decisionmaking would be helpful for a judge or jury.

Behavioral insights thus offer the possibility of enriching legal doctrine, and reducing the extent to which it is both under- and over-inclusive, with a more accurate understanding of why people behave the way they do.<sup>237</sup>

### *C. Implementing Human Rights Treaties*

The implications of behavioralism for the implementation of human rights commitments are potentially significant. Very broadly, they counsel for paying greater attention to the social mediums that were so prominent in the above examples. Media, for example, are a crucial component in the Rwanda case; the Ghana case; the Afghanistan case; and the Egyptian case. In each of these, except perhaps the Ghana case, the medium itself conveyed a message separate from the content: listening to a radio broadcast in Afghanistan, the listener imagines an entire community of fellow listeners; on Facebook's social networking software, the relationship between the content and the social situation is explicitly mapped on the page. These cases share other similarities, too: media's abilities to shape behavior stems from their effects not only on the cognition of the target actor(s) but more significantly on the actors' social situations.

Taken together, these examples have institutional implications. They suggest that social regulation can and does occur without the involvement of the state. That potentially means that some nonstate actors should be regarded as regime architects. But it also expands the regulatory toolkit of the state, and might have implications for the roles of the state's various regulatory players: shifting attention to media and education means shifting resources and priorities from certain actors (for example, the police) toward other, different actors (such as national human rights commissions, teachers, and speech writers).

This section outlines three design implications. First, the behavioral approach calls for an awareness of social situations as powerful regulators of behavior. That may, as a matter of course and depending on the regulatory goals, call for doing nothing—or rather, may counsel for deference to existing social structures and processes. Such a determination requires careful study of the social situation, for which the many tools offered by empirical insights may be useful.

Second, a behavioral approach would counsel for regulating the social medium directly. Like the human rights non-governmental organization that uses radio to rouse public sentiment in Afghanistan, human rights regimes

---

237. There are many other applications beyond criminal group behavior. Determining the limits of civil liability for international human rights abuses, for example, is another potential area for situational enrichment. In Alien Tort Statute litigation, courts have recently struggled to define the limits of aiding and abetting liability. *See, e.g., In re South African Apartheid Litigation*, 346 F. Supp. 2d 538 (S.D.N.Y. 2004).

must harness existing social structures and influence pathways. This approach may also suggest that human rights education programs reach beyond classrooms to public cinema halls and billboards.<sup>238</sup> Third and finally, such an approach would seek to expand situations that promote human rights, and contract or curtail those that, generally speaking, nurture the violation of those rights.

### 1. *Mapping Existing Networks; Stepping Aside*

Regime designers who seek to create regimes that reflect the power of social situations must carefully calibrate their regimes to existing social realities. This task necessitates an inquiry into the social substrate to find out whether or not the contemplated behavior is already being regulated by existing social situations. If it is, then the next question is how the behavior is being regulated, and whether to further regulate the behavior through direct engagement with the situation, indirect engagement through formal law, or the more conventional path—direct regulation through formal law. Such an analysis might insure against the threat that direct formal regulation would counteract or interfere with informal social mechanisms that are already doing—or could do—the necessary work.

As part of the first step of understanding social situations, a behavioral human rights regime could map existing social structures as networks. This would allow regime designers to apply the tools of social network analysis. Activists have already used network analysis to map and influence complex organizations, such as slumlords, political contribution networks, and corporate interlocks.<sup>239</sup> But conventional international legal thinking does not explicitly reflect insights from social network analysis.

Today, social networking and peer-production technologies explicitly aim to enhance a group's ability to create, maintain, and expand or contract social networks. American President Barack Obama's social networking software, my.barackobama, is a recent example, credited with enabling him to send targeted messages and to raise record sums of money.<sup>240</sup> The utility of such software is hotly debated, but most scholars agree that social software at least serves to make extant networks salient. This is significant because one of the defining characteristics of a network is its distribution, or its "scale."

---

238. For example, the Indian Supreme Court recently held that cinema halls and video parlors have a duty, through licensing, to exhibit free of cost at least two messages on the environment in each show. See *M.C. Mehta v. Union of India & Others* (1992) 1 S.C.C. 358.

239. See SKYE BENDER-DEMOILL, *POTENTIAL HUMAN RIGHTS USES OF NETWORK ANALYSIS AND MAPPING: A REPORT TO THE SCIENCE AND HUMAN RIGHTS PROGRAM OF THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE* (2008).

240. See Joshua Green, *The Amazing Money Machine*, ATLANTIC, June 2008, at 60 ("Almost overnight, a whole new network, which would yield its own spin-offs, had come into being and gone to work for the candidate.").



Networks are described as heterogeneous, or “scale free,” if there is no typical (modal) number of links for each node. Early network scientists predicted that nodes in a network would be distributed evenly, as if in a fishing net—but large social networks tend to be scale-free.<sup>241</sup> In scale-free networks, connections are distributed unequally according to a power law distribution. This distribution prevents the general features of the network, such as the average number of links per node, from being used to make predictions about the network. For example, in the case of HIV-transmission networks, some nodes may have a limited number of sexual partners, while a few nodes have hundreds of partners, meaning that attempts to predict or counteract the spread of the disease will fail if they are aimed at the modal node.<sup>242</sup>

Political scientists Lake and Wong have also suggested that in scale-free networks, centrality takes on particular significance because it can describe a node’s political opportunity structure.<sup>243</sup> The “hubs” in a scale-free network act as go-betweens for outlier nodes, and as such can control what information flows to whom and when.<sup>244</sup> We can see how this functioned in the Rwandan genocide to give power to those controlling the radio broadcasts, while limiting the power of outlier nodes and increasing their dependency on the radio.

Because political scientists are just turning to network science, there are few theories to explain how a node’s position in a network relates to its power in the network.<sup>245</sup> In one of the only studies to examine network power in the human rights context, Lake and Wong suggest that central nodes benefit from the ability to set agendas, and this agenda-setting power is especially influential in the early stages of network formation.<sup>246</sup> Hafner-Burton and Montgomery suggest that international governmental organizations have a structural power by the very fact of their centrality, which acts as a form of social capital.<sup>247</sup> Others have focused on actors’ ability to access the information in a network, which may be a function of network posi-

241. See generally GUIDO CALDARELLI, *SCALE FREE NETWORKS* (2007).

242. See Strandburg et al., *supra* note 63, at 1304 (“Any predictions of how fast the virus will spread based on the average number of sexual partners will be completely wrong. Moreover, a program of random condom distribution or HIV testing may be much less effective than anticipated if it happens to miss the few most highly active individuals.”).

243. See Lake & Wong, *supra* note 66, at 127–28.

244. But the idea that key influencers should be targeted for information campaigns is hotly debated by network scientists at the moment. See Clive Thompson, *Is the Tipping Point Toast?*, *FAST COMPANY*, Feb. 2008, at 74 (noting that Duncan Watts’ work on scale-free networks has put into doubt the claim that key nodes in a network can be identified *ex ante* and used for marketing purposes).

245. See, e.g., DAVID GREWAL, *NETWORK POWER* (2008) (decrying the lack of political theory addressing networks).

246. See Lake & Wong, *supra* note 66, at 128.

247. See Emilie M. Hafner-Burton & Alexander H. Montgomery, *Power Positions: International Organizations, Social Networks, and Conflict*, 50 *J. CONFLICT RES.* 3, 11 & n.9 (2005) (noting their reliance on “the idea that structural holes are sources of capital and the idea that centrality is a source of capital. We take the latter definition as our basis for measuring social power.”) (citations omitted).

tion.<sup>248</sup> The design implications for regime architects seem straightforward: central positions should be coveted, for reasons of leverage as well as information dissemination.

But most importantly, these insights simply suggest that social structure matters and that before any particular intervention can be made, a map of those structures and how they interact will be necessary. Once those relationships have been mapped, regime architects can decide whether to harness or change them—or do nothing. This third option is the most interesting because it is the one least often consciously pursued by practitioners. The models discussed above do not account for organic, non-intervention-based solutions to social conditions that engender human rights abuses. Network mapping may provide practitioners with the insight to know when to step in, and when to step aside.

## 2. *Harnessing Existing Social Situations*

International regulators who seek to harness existing social fabrics do not have to start anew: thick grooves have already been carved by market networks, religious networks, customary networks and professional networks. Rather than seeking state enforcement of laws that purport to protect international human rights but may have little local normative purchase, a behavioral approach to human rights practice would suggest teaming up with private corporations to harness their marketing resources, their distribution channels, and their currency (literal and figurative).

Religious networks, for example, are non-legal networks that regulate behavior through norms. A behavioral approach to human rights regime design would counsel for identifying points of synergy between human rights principles and major religious principles before seeking to embed themselves into local normative and belief structures. But conventional accounts of human rights law tend to see religion as a barrier, not an enabler, of rights.<sup>249</sup> Without laying out a strategy for how to effect this change, and without going into the normative implications in each case, by taking social situations seriously we see that an attempt to link human rights norms and religious norms may be a powerful if not necessary step. And, given the

---

248. See Strandburg et al., *supra* note 63, at 1313–14 (“Social network studies demonstrate that access to information depends on one’s position in the network. An individual’s susceptibility to legal and social sanctions for socially undesirable behavior will be similarly dependent on his or her local social network.”).

249. There are a few academics who have attempted to square human rights norms with aspects of religion. See, e.g., Abdullahi A. An-Na’im, *Islamic Foundations of Religious Human Rights*, in *RELIGIOUS HUMAN RIGHTS IN GLOBAL PERSPECTIVE: RELIGIOUS PERSPECTIVES* 337 (John Witte, Jr. & Johan D. van der Vyver eds., 1996). See also Barbara Ann Rieffer, *Religion, Politics and Human Rights: Understanding the Role of Christianity in the Promotion of Human Rights*, 6 *HUM. RTS. & HUM. WELFARE* 31 (2006), available at [www.du.edu/korbel/hrhw/volumes/2006/rieffe-2006.pdf](http://www.du.edu/korbel/hrhw/volumes/2006/rieffe-2006.pdf). Still, these attempts to reconcile religion with human rights principles remain a tiny percentage of the scholarship.

problem of limited resources endemic to human rights work, many religions offer a ready-made normative project with which to work.

This may mean that human rights regimes, like individuals in society, must embrace ambiguity and exhibit somewhat contradictory behaviors. For example, the behavioral human rights regime would seek to harness corporations for pro-social ends, even while in other contexts seeking to regulate the corporation itself. Corporations can be grave rights abusers—but rather than asking a developing state to sanction a corporation that brings business and jobs to the state, human rights practitioners could aim to embed their values into the brand and institutional culture of the corporation itself. This is precisely the idea behind several of the most innovative brand strategies pursued by human rights lawyers.<sup>250</sup> And this could happen while practitioners are seeking to regulate the behavior of key nodes in the corporate governance structure.

This sort of regime may also require state actors to act in ways conventionally not considered part of law-making. Just as police work in American cities has come to include community relations and anti-graffiti programs, and military operations have come to include school building and other population-based strategies, so too international legal regime builders may have to focus on social structures in conjunction with—and in some cases as antecedents to—their formal legal structures.

Attention to human rights situations might suggest a shift away from calls to get police to enforce explicit violations of rights, and toward a closer focus on creating rights-positive situations. This is already happening in some countries, under the rubric of “human rights education,” which has been defined as “a deliberate, participatory practice aimed at empowering individuals, groups and communities through fostering knowledge, skills and attitudes consistent with internationally recognized human rights principles.”<sup>251</sup> While classroom education is an essential component of this, it is only a piece of the larger project. Most states have an administrative body—a pre-existing network—known as a National Human Rights Institution.<sup>252</sup> These bodies are charged with monitoring state compliance with human rights obligations, as well as engaging in rights education. The U.N. General Assembly described human rights education as “a comprehensive life-long process by which people at all levels of development and in all strata of society learn respect for the dignity of others and the means and methods of ensuring that respect within a democratic society.”<sup>253</sup> That could be read to

---

250. See Lance Compa & Tashia Hinchliffe-Darricarrère, *Enforcing International Labor Rights Through Corporate Codes of Conduct*, 33 COLUM. J. TRANSNAT'L L. 663 (1995).

251. See, e.g., Amnesty International, *What is Human Rights Education?*, <http://amnestynepal.org/content/hre.php>. (last visited Oct. 31, 2009).

252. As of the writing, over 100 states have a national human rights institution. See *The National Human Rights Institution Forum*, available at <http://www.nhri.net/nationaldatalist.asp>. (last visited Oct. 30, 2009).

253. G.A. Res. 48/127, ¶ 5, U.N. Doc. A/RES/48/127 (Dec. 20, 1993).

be vague, undirected moralizing—but it could also be read as a call to harness social situations for human rights.

This aspect of the human rights regime is conventionally overlooked as either weak or insufficiently legal because it does not match the strict law enforcement model discussed above. But successful human rights education is a necessary pre-requisite for the rule of law. Children's education, adult education, professional education, and media campaigns are all "cheap talk" to a rational actor, and may therefore come cheap in treaty negotiations with recalcitrant states—but they are potentially powerful tools in the social medium. That makes them a good value.<sup>254</sup> Studies of human rights education programs suggest that they very cheaply and effectively change peoples' attitudes toward human rights (for the better) and correspondingly change behavior.<sup>255</sup>

In many cases, the behavioral approach to human rights will not be revolutionary—for example, human rights regimes already contain provisions for creating a human rights culture.<sup>256</sup> Likewise, the importance of informational campaigns has long been known to human rights practitioners. But behavioral insights may shift the focus of human rights regimes to think more tactically about implementing human rights education and human rights media. Two possible innovations include outside monitoring of human rights education; and, as noted above in the commitments recommendation, outside productive and dissemination of human rights education and media, bypassing the state. For those countries that have populations with regular access to the internet, but that might not have been conducting effective human rights education due to capacity constraints, new online media may prove powerful as well.

International legislators seeking to regulate social situations will find many benefits in partnerships with these influential networks. These networks are able to not only identify influence pathways but also often serve as the carriers of the relevant international legal norms. They can also be enlisted in partnerships that lend credence to the norm project, allowing the international norm advocate to become a culturally relevant player in the social space. This sort of regime would feature public-private partnerships and would identify and utilize preexisting and nascent influence pathways. In addition to mapping and understanding existing institutions and influ-

---

254. See Dickerson, *supra* note 212, at 567 ("[A] targeted focus on a network's richest nodes can be the cheapest way to destroy—or influence—the institution, and thus to guide the organization's actions.")

255. See J. Christopher Cohrs et al., *Determinants of Human Rights Attitudes and Behavior: A Comparison and Integration of Psychological Perspectives*, 28 *POL. PSYCHOL.* 441 (2007) (using a time-lagged internet survey (N=479) to analyze individual orientations to human rights, and finding that political ideology as well as human rights knowledge predicted human rights behavior).

256. See, e.g., United Nations Convention on the Elimination of All Forms of Discrimination Against Women, Preamble, Sept. 3, 1981, 1249 *U.N.T.S.* 13 ("Aware that a change in the traditional role of men as well as the role of women in society and in the family is needed to achieve full equality between men and women.").

ence pathways, situational regulators might then deepen those grooves, and strengthen particular networks, or nodes within a particular network, in order to achieve the desired regulation. New tools—such as cell phones—make this work much easier because they not only make networks salient but also enable a group to reach many people in the network at once.

### 3. *Changing the Social Situation*

In addition to harnessing existing pathways for social regulation, regime designers can create new spaces for social influence, and create—or eliminate—new grooves in the social substrate. We have already seen that news media can be important existing influence pathways, and as such they play an important role in determining the robustness of the social regulatory space. In some contexts—as in the Rwandan genocide—the paucity of media outlets and the limited access many people had to alternative news sources gave the broadcasters enormous behavioral-regulatory power. In such situations, increasing the number of available media outlets serves to diffuse that power. For example, Interactive Radio for Justice allows people in different parts of Ituri to report on news from their otherwise isolated war-stricken area by text message or phone call.<sup>257</sup> The information pathway did not exist before the network was established—and once it was in place, the text message had the potential power of a news broadcast.

Unlike conventional media, new media are usually two-way, or “many-to-many.”<sup>258</sup> Examining network analysis insights about the difference between one-way and two-way transmissions, it quickly becomes apparent that the difference between a radio and a networked computer is not just one of degree. The power of these technologies is not just that they allow one user to broadcast information, but rather that they create, in some instances, spaces for dialogue and identity formation. New two-way technologies—cell phones and internet-based communications like email and discussion boards—are important because they create the virtual version of physical spaces for dialogue.<sup>259</sup>

New spaces for dialogue mean new opportunities for norm-internalization. Dialogue and internalization are situationally dependent. They are also crucial for achieving consensus.<sup>260</sup> Behavioral regulators, then, might seek to

---

257. See Stephanie Hanes, *Backstory: Congolese Radio Show Gives War Victims a Voice*, THE CHRISTIAN SCI. MONITOR, Mar. 22, 2007, at 20.

258. See CLAY SHIRKY, *HERE COMES EVERYBODY* 87 (2008).

259. This Article is not sanguine about the use of technology to challenge state powers. See, e.g., GOLDSMITH & WU, *supra* note 207. As the Chinese government's internet censorship efforts have shown, certain internet messages and pages can be effectively censored by a state. But as the Egyptian example shows, even states that heavily censor their print and broadcast media may not currently have the resources or knowhow to censor online material. Likewise, we might imagine that some states will choose to allow certain limited expansions of individual freedom, such as the ability to assemble on Facebook.com, in order to reap financial or other gains as long as the state sees the gains as greater than the costs.

260. See BICCHIERI, *supra* note 20, at 167–70 (noting that perceived consensus increases after discussion—and that this was even true when subjects were discussing with inanimate computers). *But see* CASS

maximize spaces where new uptake of ideas and norms occur. In the human rights context, this could mean holding forums where stakeholders in an issue convene to discuss their shared experiences. Regime designers might seek to create similar spaces to ensure uptake of norms and interaction among certain nodes in a network—the creation of such spaces can enhance or interrupt static network dynamics.

These spaces also provide users with new avenues—some potentially dangerous—for social influence. For example, the director of Facebook's Causes—a subset of tools for users of the social network who wish to promote certain social and political causes—has shown that when one user in a network selects a particular social cause, and that change is 'broadcast' throughout the network, there is a higher likelihood that someone else in the user's network will also take up that particular cause.<sup>261</sup> These networking tools can also be used for coordinated action based around a particular network norm. In wealthier countries, where most of the funding for human rights work in poorer countries originates, social networking tools are already proving to be powerful fundraising and social mobilizing tools.<sup>262</sup> But exposed networks mean that actors seeking to violate rights also have new-found social power.

Behavioralism helps here, too, for it offers insights about how to design the limits on situations that encourage rights violations. The ICTR—the same tribunal that decided the *Nahimana* case—recently issued a judgment against Simon Bikindi, a Hutu musician who had written and circulated songs that claimed the superiority of Hutus over Tutsis. Bikindi was convicted, however, for a single speech act that loudly and directly called for Hutus to kill Tutsis.<sup>263</sup> He was not convicted for his huge role as a cultural voice for the genocidaires—nor as a norm entrepreneur among Rwandans.

A behavioral approach to Bikindi's actions would have involved a much more nuanced analysis of his position in the social fabric of Rwanda during the preparations for and execution of the genocide. It would have examined the role his narratives played in constructing the identities of Hutus who later slaughtered Tutsis in what could be seen as the result of "channel factors"—the following of a social script that Bikindi deliberately sought to create through custom and ritual. Music, we know, is a powerful tool for creating and sustaining social movements.<sup>264</sup> In his study of ritual and the

---

SUNSTEIN, GOING TO EXTREMES: HOW LIKE MINDS UNITE AND DIVIDE 1–5 (2009) (noting that evidence suggests that groups of like-minded people can radicalize rather than moderate one another's views).

261. Interview with Joe Green, Founder, Facebook Causes (Dec. 2, 2008) (on file with the Harvard Law Library).

262. See Green, *supra* note 240, at 60.

263. See *Prosecutor v. Bikindi*, Case No. ICTR 01-72-T, Judgment (Dec. 2, 2008).

264. See RON EYERMAN & ANDREW JAMISON, MUSIC AND SOCIAL MOVEMENTS: MOBILIZING TRADITIONS IN THE TWENTIETH CENTURY (1998).

law, Andrew Cappel analyzes the importance of rituals for the uptake and internalization of norms.<sup>265</sup>

The court's decision did not make use of social science insights to explain Bikindi's role as a social regulator, and in not doing so missed an important opportunity. There is a role for social situational insights in designing and tailoring regulations and in their interpretation by a court in order to limit certain rights-threatening situations. Behavioral insights suggest that the power of media derives from their ability to drive certain narratives and the corollary identities and norm sets. International legal regime designers, therefore, must pay attention not just to the scope of media outlets, but also to the narratives they broadcast and whether they will "stick."<sup>266</sup>

#### IV. TOWARD A BEHAVIORAL MODEL

The foregoing, while not a complete sketch of a behavioral regime, is enough to illuminate some of the benefits and limitations of this approach.<sup>267</sup> Behavioralism offers tools for seeing and mapping existing social structures, which widen the scope of regulatory options; it introduces new roles for actors not traditionally considered significant contributors to international law; it offers a new source of tools for emerging conceptions of global governance; it is concerned with shaping behavior *ex ante* as much or more than with enacting *ex post* retribution for behavior; and finally, it offers a new approach to old debates. But behavioralism raises concerns about both its utility and whether it is compatible with liberalism.

##### A. Advantages

Behavioralism offers new avenues for human rights regime design. It does not *per se* make old avenues obsolete, but in some cases it suggests a shift in priorities, including: realizing the importance of economic, social and cul-

---

265. See Andrew J. Cappel, *Bringing Cultural Practice into the Law: Ritual and Social Norms Jurisprudence*, 43 SANTA CLARA L. REV. 389, 485 (2003). Likewise, attempts to shape behavior can be undermined if they do not take account of potentially counter-productive customs. In many refugee settlements, for example, customary forms of dispute settlement must be merged with U.N.-sanctioned legal dispute resolution mechanisms—for when they are ignored, they operate as a secondary system undermining the official, U.N.-sanctioned one. See U.N. High Comm'r for Refugees, *The Administration of Justice in Refugee Camps: A Study of Practice*, U.N. Doc. PPLA/2006/01 (Mar. 2006) (prepared by Rosa da Costa), available at <http://www.unhcr.org/refworld/docid/4417f9a24.html> (examining refugee dispute settlement mechanisms through surveys and site visits to 52 camps in 13 countries with a total of nearly a million refugees).

266. See CHIP HEATH & DAN HEATH, *MADE TO STICK* (2007).

267. This Article does not address the foundational challenges to human rights, which are beyond the scope of its inquiry. To some, the behavioral approach cannot be squared with rights-talk. This Article agrees with Richard Rorty when he says that "[t]he best, and probably the only, argument for putting [human rights] foundationalism behind us is the one I have already suggested: It would be more efficient to do so, because it would let us concentrate our energies on manipulating sentiments, on sentimental education." Richard Rorty, *Human Rights, Rationality and Sentimentality*, in *ON HUMAN RIGHTS: THE 1993 OXFORD AMNESTY LECTURES* 112, 122 (Susan Hurley & Stephen Shute eds., 1993).

tural rights that have traditionally taken a back seat to civil and political rights; maximizing the role of education and media as cultural and therefore regulatory tools; and focusing on mechanisms that the fundamental attribution error suggests will be more powerful than regulators might estimate when they evaluate or establish them.

Because this approach depends on social situations for its regulatory effect, it does not entirely depend on powers that are unique to the state. This is another benefit of the behavioral approach: it fixes our gaze on actors who are ignored or under-appreciated in other accounts of international legal enforcement. Nonstate actors are, with some exceptions,<sup>268</sup> under-represented as agents of change in the international legal scholarship, while in fact they are poised to have a significant impact in practice, given their often deeply embedded position in social substrates where the translation occurs in so many legal models between international and local norms.<sup>269</sup>

The behavioral approach is also explicitly concerned with ex ante norm promotion in addition to the traditional concern with ex post sanctions for rights abuses. One goal of ex post sanctions is to send a message to potential abusers about the risks of their actions.<sup>270</sup> But this deterrence effect only works if the targeted actor recognizes and can internalize the threat of a sanction; behavioral insights suggest that in fact that sanction may be competing with many stronger social forces for influence over the actor's behavior. By thinking about how norms can be activated in a particular situation, the behavioralist approach returns our focus to the norms themselves rather than the sanctions that purport—rightly or not—to advance those norms. This approach offers an alternative, as yet underappreciated avenue for scholarship and norm promotion. That might also mean, in certain cases, doing nothing. At the very least, behavioral insights suggest a more careful examination of social substrates that might help policy makers to decide when not to act given that heavy-handed rules might crowd out more effective social regulatory process.

But these are tactical, not moral decisions. That is, such insights do not necessarily threaten existing moral conceptions of justice—rather, they sug-

---

268. Political scientists Margaret Keck and Kathryn Sikkink offer convincing and detailed explanations for how nonstate actors wield power internationally. But their work has not been updated with empirical research about the micro-processes of influence. Where they describe the persuasive value of issue framing, for example, they say little to nothing about how nonstate actors design or interact with social situations to enact change. See KECK & SIKKINK, *supra* note 38. Similarly, Anne Marie Slaughter's work made the notion of governance networks salient, but her work drew little from network science, and was limited to networks of government officials, limiting the explanatory action of her model to formal regulators. As a consequence, she focuses on persuasion, ignoring other potentially more important forms of influence such as network position and/or acculturation. See ANNE MARIE SLAUGHTER, *A NEW WORLD ORDER* (2004).

269. See, e.g., Koh, *supra* note 140, at 207. His account includes nonstate actors as bodies who transmit global norms to local communities, but their own regulatory systems are under-explained, as is their ability to form horizontal global regulatory regimes without state involvement.

270. See, e.g., Paul H. Robinson & John M. Darley, *The Role of Deterrence in the Formulation of Criminal Law Rules: At Its Worst When Doing Its Best*, 91 GEO. L.J. 949, 956 (2003).



gest a rethinking of existing mechanisms for delivering justice. Recalling the Rwandan genocide, Kwame Anthony Appiah explains:

The conjunction of virtue ethics and situationism urges us to make it easier both to avoid *doing* what murderers *do* and to avoid *being* what murderers *are*. We can admire moral heroism while deploring the circumstances in which it can arise. It's good to feel compassion; it's better to have no cause to.<sup>271</sup>

Behavioralism, then, may also offer a way around some of the most entrenched debates in international legal scholarship—not only those noted above dealing with state compliance, but also cultural critiques that charge rights regimes with pandering to immoral behaviors in the name of sensitivity. As Madhavi Sunder wrote in disgust about religious fundamentalists in Afghanistan who violated what she called the “fundamental human rights” of women: “[L]egal norms such as the ‘freedom of religion,’ the ‘right to culture,’ and the guarantee of ‘self-determination’ defer to the claims of patriarchal, religious elites, buttressing their power over the claims of modernizers.”<sup>272</sup> Sunder offers one approach to the human rights community’s trepidation about cultural relativism: get over it. The behavioral approach offers another: get around it.

Certainly culture has the power to create or undermine human rights norms. In Sunder’s account, Afghani religious beliefs (culture A) conflict with human rights (culture B). But rather than taking an entire set of cultural values head-on in this manner, the behavioral approach focuses on the social conditions that enable or disable certain actions. This approach circumvents questions of rights essentialism entirely because the central focus is on behaviors, not values. Cultural values only enter the discussion as pieces of a larger social situation.<sup>273</sup>

Finally, perhaps most importantly, the behavioral approach may have significant implications for other areas of international law, especially the nascent attempts to understand global governance in a time of great global change. There are at least three different ways in which the behavioralist approach to human rights regimes offered here operates in service of the law. In the first instance, we might imagine that there is a not insignificant set of states that want to enforce international human rights law—or take on new international commitments—but do not do so because of capacity constraints or fears about capacity constraints.<sup>274</sup> Ghana is one of these states, eager to uphold its international commitment to the human right to health but unsure about how to do so. Behavioral insights help Ghanaian health

---

271. APPIAH, *supra* note 194, at 72.

272. Madhavi Sunder, *Piercing the Veil*, 112 YALE L. J. 1399, 1403 (2003).

273. Of course, behaviors are associated with values—and the behavioral approach will not resolve cultural relativist charges. But as a tactical matter, it may be more productive for human rights practitioners and scholars to focus on—and speak in terms of—behaviors rather than norms or cultural values.

274. See CHAYES & CHAYES, *supra* note 190, at 10.

officials, and might help other states as well, to actualize norms and rights that lie dormant in unproductive social situations. For example, a state might hesitate to meaningfully implement the International Covenant on Economic, Social and Cultural Rights for fear of calls to redistribute private wealth in the name of the human right to health. A behavioral approach to that right, however, might imply a different set of commitments by the state that ultimately amounts to a lower threshold for ratification and a higher likelihood of enforcement.

Second, the behavioral approach serves to strengthen existing international law and legal enforcement mechanisms. For example, behavioral insights may help to make international obligations more binding. Commitment strategies and specifically designed social situations that harness human biases may result in more net compliance with international law.<sup>275</sup> Or consider an adjudicatory body using behavioral insights to make sense—or rather, better doctrine—out of vague international human rights obligations, as the ICTR could have done in the *Nahimana* and *Bikindi* cases. Legal doctrines evolve, and so it would be easy to imagine, say, joint criminal enterprise liability coming to reflect empirical insights about the power of social situations. It is also true that, through what is known as the expressive function, laws express norms in an attempt to create or curb them.<sup>276</sup> The behavioral approach would augur a shift in regime priorities from the mechanisms of enforcing the norms embodied in human rights law to the very situations in which human rights law has an expressive function—that is, where a norm is activated. Just as Cass Sunstein has written that “norm management is an important strategy for accomplishing the objectives of law, whatever those objectives may be,”<sup>277</sup> so too we might say that social dynamics are important for accomplishing the objectives of international law.

Third, and finally, behavioral insights help regime designers to create social situations—again, Cover’s “normative universes”—where human rights norms have social meaning.<sup>278</sup> This meaning-making can be said to be a legal production—in Cover’s words, a “jurisgenerative” product.<sup>279</sup> This is the most theoretical piece of the claim but it may also be the most important. We say that computer code acts as a regulatory regime insofar as it is a text that dictates the limits of one’s behavior online.<sup>280</sup> We say that custom and community determine social norms, which in turn govern behavior in a given context.<sup>281</sup> And we say that multinational corporations create informal regulatory regimes through their contracts, market influence, and marketing

---

275. See *supra* Part III.A.

276. See Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L. REV. 903, 964–65 (1996).

277. *Id.* at 907.

278. See Cover, *supra* note 25, at 7.

279. *Id.* at 15.

280. See LESSIG, *supra* note 139, at 5–6.

281. See ELLICKSON, *supra* note 36.

ability.<sup>282</sup> Behavioral insights suggest that we can also say that social situations act as regulatory mediums—places for normative influence. In this way, behavioralism is a project that dovetails with, and may prove useful to, nascent scholarship on global governance.

As administrative law scholar Jody Freeman notes, regulation is more a question of governance than of governments, where governance is a “set of negotiated relationships between public and private actors.”<sup>283</sup> Numerous global governance scholars have produced many meta-theories to explain global regimes that fall below the level of formal treaties between states, including: global administrative law;<sup>284</sup> constitutionalism;<sup>285</sup> bottom-up legal process, or the New New Haven School.<sup>286</sup> But these projects are in their infancy, and they rarely make use of empirical support to explain the social processes of influence in the “set of negotiated relationships”<sup>287</sup> that drive global governance. As David Kennedy notes, these uncharted relationships remain today “the mystery of global governance.”<sup>288</sup> Critiquing current scholarship for ignoring such social structures, he warns: “We have paid too little attention to the myriad ways power flows through the capillaries of social life, perhaps particularly at the global level.”<sup>289</sup>

Kennedy surveys the existing global governance literature, noting that the most exciting and uncharted frontier is the social domain often thought to exist outside of law. The scholars, he says, “pay too little attention to the informal and the clandestine—to customary norms, background patterns of private and public expectation, black markets and illegal flows,” all of which are potentially potent sources of global governance.<sup>290</sup> As Kennedy’s charge implies, focusing on social situations offers a pathway to the ground-level questions about how social processes relate to larger legal-regulatory systems, and it allows us to identify potential design levers for public policy at the domestic and local level.

282. See Larry Catá Backer, *Economic Globalization and the Rise of Efficient Systems of Global Private Law Making: Wal-Mart as Global Legislator*, 39 CONN. L. REV. 1739 (2007).

283. Christian Joerges, *Transnational Governance and its Legitimacy Problematics: The Examples of Standardization and Food Safety*, 7 (Feb. 2, 2004) (unpublished paper) (citing Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. REV. 543 (2000)).

284. See Benedict Kingsbury, Nico Krisch & Richard B. Stewart, *The Emergence of Global Administrative Law*, LAW & CONTEMP. PROBS., Summer/Autumn 2005, at 15, 46.

285. See Gunther Teubner, *Societal Constitutionalism: Alternatives to State-Centered Constitutional Theory?*, in TRANSNATIONAL GOVERNANCE AND CONSTITUTIONALISM 3 (Christian Joerges et al. eds., 2004).

286. There is a movement some have labeled the New New Haven School, a riff on the eponymous international relations school of thought. See, e.g., Janet Koven Levit, *A Bottom-Up Approach to International Lawmaking: The Tale of Three Trade Finance Instruments*, 30 YALE J. INT’L L. 125 (2005). See also Melissa A. Waters, *Normativity in the “New” Schools: Assessing the Legitimacy of International Legal Norms Created by Domestic Courts*, 32 YALE J. INT’L L. 455 (2007).

287. Joerges, *supra* note 283.

288. David Kennedy, *The Mystery of Global Governance*, 34 OHIO N.U. L. REV. 827 (2008).

289. *Id.* at 851.

290. *Id.* at 849.

The Global Administrative Law scholars acknowledge the limitations of drawing on domestic administrative law regimes—not because they offer too radical an approach, but because they might exclude important or emerging forms of transnational regulation.<sup>291</sup> This inquiry into the sources of international law is fundamental:

It may be that a better account of the legal sources of existing normative practice in global administration could be grounded in a revived version of *ius gentium* that could encompass norms emerging among a wide variety of actors and in very diverse settings, rather than depending on a *ius inter gentes* built upon agreements among states. This approach would mirror, to some extent, law-making procedures in other fields of law beyond the state, such as the *lex mercatoria*, based on the practices of commercial actors worldwide.<sup>292</sup>

The behavioral approach at least attempts to explain the processes through which local situations activate certain norms. If international legal scholarship makes the theoretical turn it appears poised to, toward *ius gentium*, empirical insights about the power of social situations may prove invaluable.

The German sociologist Gunther Teubner has theorized even broader claims about transnational social regimes—what he calls global constitutionalism. Law-making, he says, “also takes place outside the classical sources of international law, in agreements between global players, in private market regulation by multi-national concerns, internal regulations of international organisations, inter-organisational negotiating systems, worldwide standardisation processes that come about partly in markets, and partly in processes of negotiation among organisations.”<sup>293</sup> Teubner advances the field of inquiry by considering the consequences of a broader understanding of the sources of international law. The problem is devising an appropriate theoretical framework for an international order no longer dominated by nation-states alone.<sup>294</sup> The behavioral approach offers a starting point for these questions.

There is a sign in front of the Perse Girls School in Cambridge, England, that reads: “Parking here is both inconsiderate and illegal.” The Perse parking police find it useful to employ multiple regulatory regimes: morality, financial incentives, fear, shame, *inter alia*. These are not just competing regulatory systems but rather complementary and deeply inter-related ones. We say of a person that she has a strong or weak, good or wicked constitution—but we stop short of investigating the relationship between her constitution and constitutionalism. Much has been written about the United

---

291. See Kingsbury et al., *supra* note 284, at 28.

292. *Id.* at 29.

293. TEUBNER, *supra* note 285, at 16 (citation omitted).

294. *Id.* at 4.

Kingdom's unwritten constitution<sup>295</sup>—and yet very little about its relationship to British behavior. This gap is a significant theoretical challenge for international legal scholars to bridge. Behavioral insights may help.

### B. Limitations

The behavioralist approach is not without its limitations. Design applications for regulating or promoting certain social situations are, as Lessig notes, “totalizing,” in the sense that they tempt totalitarian abuse.<sup>296</sup> One of the chief objections to the behavioral approach—which may call for manipulation of the social situation, as the Ghanaian handwashing example illustrates—is that of paternalism. Behavioral economists and legal scholars drawing from behavioral insights have been extremely sensitive to this charge. Richard Thaler and Cass Sunstein explicitly named their theory for how certain behavioral insights can improve public policy “libertarian paternalism” in a bid to preempt paternalist critiques.<sup>297</sup> Libertarian theorists like Charles Fried have proposed theories of liberty that are at odds with situation and norm management, which are reviled because they are seen to constrain choice.<sup>298</sup>

But built into the very concept of paternalism is an assumption about human agency—namely that it is synonymous with liberty as the absence of constraint. The behavioral approach challenges this very notion, suggesting that all social situations involve some form of constraint, and that a more accurate description of the libertarian move would be an absence of perceived constraint. If we acknowledge that situational factors are all around us, and that they act as unperceived influences (constraints) on behavior, then the paternalistic charge becomes less persuasive. We are left, instead, with a decision about which sorts of constraints to allow and which to curtail. Liberty, according to this view, is more a function of one's ability to make informed choices aware of the relevant social influences.

If we take the behavioral charge seriously, concerns over paternalism are just another form of what psychologists term the act-omission bias—making a distinction between a willful act and a willful omission that have functionally the exact same result.<sup>299</sup> As Thaler and Sunstein suggest, someone has to make choices and they should do so carefully—that is, with all

---

295. See, e.g., ANTHONY KING, *THE BRITISH CONSTITUTION* (2007) (arguing that the famed social fabric constitution has actually been transcribed in the form of thousands of pages of laws and regulations that govern daily life in Britain).

296. See Lessig, *The New Chicago School*, *supra* note 24, at 691.

297. See CASS SUNSTEIN & RICHARD THALER, *NUDGE* 4–6 (2008).

298. See CHARLES FRIED, *MODERN LIBERTY* 100 (2006) (finding liberty at odds with the idea of choice as a social construction: “If we do not take responsibility for our own judgments, we might as well give up on thought and argument . . .”).

299. See JONATHAN BARON, *THINKING AND DECIDING* 504–05 (3d ed. 2000) (“The ‘omission bias’ is the bias that favors harms of omission over equivalent harms from action.”)

the relevant behavioral insights at their disposal.<sup>300</sup> A British research group recently conducted a study that found similar sentiments among British residents. The study revealed that the majority of British residents want the government to intervene in their lives in order to protect them, but that they also think the government does not adequately trust them to make important decisions for themselves regarding dangerous activities.<sup>301</sup> A related study shows that in Europe, the size of government social programs is positively correlated with the “vibrancy of the civic sphere,” which suggests high levels of individual empowerment.<sup>302</sup> The idea, these studies claim, is that what policies some perceive as threats to liberty may also be seen as preferable gains in agency.

Bicchieri categorizes this problem as one between “awareness and choice,” noting that real choice is significantly influenced by the relevant situation and not merely determined by norms.<sup>303</sup> Thaler and Sunstein offer a salve for concerns about paternalism that tracks this link between awareness and choice: they suggest that any state manipulation of the social space come with absolute transparency.<sup>304</sup> But here we reach perhaps the biggest functional limitation of the behavioral approach: reactance.

Reactance is the urge to do the opposite of whatever someone is persuading you to do, if only to assert your freedom.<sup>305</sup> Psychologists have shown this to be a relatively robust phenomenon. This suggests that government attempts to promote human rights situations openly and transparently may, in fact, cause resistance to those programs. If citizens are aware of the aims of the government campaigns, they may turn away and adopt more rights-negative norms. Behavioral regime designers will have to account for reactance and shape their policies accordingly. But reactance can also be harnessed to create a situation of rule compliance. The Hawthorne Effect, for example, suggests that people perform differently when they think they are being watched, even when they are not being watched.<sup>306</sup> Scholars of prison systems have shown that oversight, or the mere perception of it, can have a significant impact on behavior.<sup>307</sup>

---

300. SUNSTEIN & THALER, *supra* note 297, at 10.

301. THE POLITICS OF PUBLIC BEHAVIOR 37 (Duncan O’Leary ed., 2008), available at [http://www.demos.co.uk/files/Politics\\_of\\_Public\\_Behaviour\\_web.pdf](http://www.demos.co.uk/files/Politics_of_Public_Behaviour_web.pdf).

302. PAUL SKIDMORE & KIRSTEN BOUND, THE EVERYDAY DEMOCRACY INDEX 72–73 (2008), available at [http://www.demos.co.uk/files/EDI\\_all%20chapters.pdf](http://www.demos.co.uk/files/EDI_all%20chapters.pdf).

303. BICCHIERI, *supra* note 20, at 46.

304. SUNSTEIN & THALER, *supra* note 297.

305. See Joseph E. Grandpre et al., *Adolescent Reactance and Anti-Smoking Campaigns: A Theoretical Approach*, 15 HEALTH COMM. 349 (2003).

306. See RICHARD GILLESPIE, MANUFACTURING KNOWLEDGE: A HISTORY OF THE HAWTHORNE EXPERIMENTS 2 (1991) (quoting F.J. ROETHLISBERGER, MANAGEMENT AND MORALE 15 (1941)).

307. See, e.g., Stanley Milgram, *Behavioral Study of Obedience*, J. ABNORMAL & SOC. PSYCHOL. 67, 371–78 (1963) (demonstrating a subject’s willingness to obey experimenter’s requests, even when those requests caused harm to others); Stanley Milgram, *OBEDIENCE TO AUTHORITY: AN EXPERIMENTAL VIEW* (1974) (expanding on the original study).

There are doubtless many other objections to the behavioral approach, but these are the most prominent and the most troubling. Still, they are ultimately not as troubling as continuing to construct and operate legal regimes with outdated information about human behavior. Behavioral insights show us that systems of control are everywhere—the relevant question is how to give the situation particular meaning. It would be truly ironic if liberal fears kept human rights regimes from constructing social situations in which human rights have meaning.

It is also true that social situations are not all-powerful. There will be cases in which actors cannot be socialized into a particular behavior and in those cases conventional regulatory tools remain helpful, indeed essential. But acknowledging the limitations of social dynamics as a tool for regulation should also assuage fears over its resemblance to all-powerful thought control.

#### CONCLUSION

Martti Koskenniemi has described the difficult task international lawyers have of steering between dichotomous poles: realism and formalism, or what he calls apology and utopia.<sup>308</sup> The aim of this Article has been to begin an investigation into whether our social constructions can help us to more effectively thread the needle between a different but related set of poles in the human rights context—those Michael Ignatieff similarly described as the twin challenges of politics and idolatry.<sup>309</sup> This Article does not offer a comprehensive answer for how to do so, but it can show in one small way how the current approach does not ask all of the right questions.

At the root of all regulatory regimes is an assumption about human behavior. Human rights regimes are no different. New information is available about the role of social situations as distinct components of human behavior. But scholars and human rights regime designers have paid insufficient attention to this new information. This Article is an attempt to expose this gap and take steps to narrow it. This is not a complete sketch, but a research agenda for work still to come and a call to invite more scholarship in this area. The first step to a comprehensive regime roadmap is to ask the right questions—questions the current approach has not asked.

The Judgment at Nuremberg set the tone for the twentieth century international human rights regime when it declared: “Crimes against international law are committed by men, not by abstract entities, and only by punishing individuals who commit such crimes can the provisions of inter-

---

308. See MARTTI KOSKENNIEMI, *FROM APOLOGY TO UTOPIA: THE STRUCTURE OF INTERNATIONAL LEGAL ARGUMENT* 565 (1989).

309. See MICHAEL IGNATIEFF, *HUMAN RIGHTS AS POLITICS AND IDOLATRY* (2003).

national law be enforced.”<sup>310</sup> But if we acknowledge that human behavior is a function of social conditions—those once-abstract entities—then a twenty-first century international human rights regime may adopt a different, more behavioral, tone.

---

310. International Military Tribunal (Nuremberg), Judgment and Sentences, Oct. 1, 1946, *reprinted in* 41 AM. J. INT'L L. 172, 221 (1947).