

1-2007

## A Theory of Expressive International Law

Alex Geisinger

Michael A. Stein

Follow this and additional works at: <https://scholarship.law.vanderbilt.edu/vlr>



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

---

### Recommended Citation

Alex Geisinger and Michael A. Stein, A Theory of Expressive International Law, 60 *Vanderbilt Law Review* 75 (2019)

Available at: <https://scholarship.law.vanderbilt.edu/vlr/vol60/iss1/2>

This Article is brought to you for free and open access by Scholarship@Vanderbilt Law. It has been accepted for inclusion in Vanderbilt Law Review by an authorized editor of Scholarship@Vanderbilt Law. For more information, please contact [mark.j.williams@vanderbilt.edu](mailto:mark.j.williams@vanderbilt.edu).

# A Theory of Expressive International Law

---

Alex Geisinger

Michael Ashley Stein

60 Vand. L. Rev. 77 (2007)

*Ever since Grotius first suggested that desire for esteem from the broader global community motivates States to comply with international law, identifying just how this desire effects compliance has proven illusive. The ability to harness the pull of international society is important to virtually all treaty formation and compliance. It is especially important in the area of human rights regimes where other compliance forces such as coercion, are rarely, if ever, used.*

*Recent empirical evidence, however, suggests that human rights regimes are ineffective. Indeed, in many situations this evidence suggests that the human rights practices of States that ratify such treaties may actually worsen after ratification. The need to understand how, or whether, the pull of international society influences state behavior, thus, has never been greater.*

*This Article provides an initial detailed model of the forces motivating human rights treaty creation and compliance by drawing on evolving expressive law literature. It begins by setting forth a need-reinforcement model that explains how normative pressure influences rational actors to alter their behavior and beliefs while seeking regard from other group members. Next, the Article applies this model to State treaty ratification and compliance, and describes how treaties exert expressive effects that lead rational States to change their behavior because of their desire to be part of and esteemed by the global community. The Article then demonstrates how this expressive theory harmonizes the contributions of divergent international law scholars into a more complete theory of why States enter into and obey international law. In doing so, it provides a framework from which regime design implications can be drawn.*

\*\*\*

# A Theory of Expressive International Law

*Alex Geisinger\* & Michael Ashley Stein\*\**

INTRODUCTION .....	78
I. EXPRESSIVE THEORY .....	83
A. <i>The Reasoned-Action Model</i> .....	83
B. <i>Beliefs as the Building Blocks of Attitude</i> .....	86
C. <i>The Need-Reinforcement Principle</i> .....	88
II. AN EXPRESSIVE MODEL OF RATIFICATION AND COMPLIANCE .....	97
A. <i>International Processes</i> .....	97
B. <i>The Interaction of International and         Domestic Processes</i> .....	99
1. Direct Influences on the Subjective Norm ...	101
2. Indirect Influences on Normative Belief.....	106
III. CLARIFICATIONS AND IMPLICATIONS OF THE EXPRESSIVE MODEL .....	109
A. <i>Harmonizing Rationalism and Constructivism</i> .....	109
B. <i>Distinguishing Persuasion, Reputation,         and Coercion</i> .....	114
C. <i>Explaining Internalization</i> .....	116
D. <i>Explaining Signaling</i> .....	118
E. <i>Implications</i> .....	120
F. <i>Making Human Rights Treaties Matter</i> .....	126
CONCLUSION.....	130

## INTRODUCTION

What is the “pull of international society” and how does it influence the willingness of States to enter into or comply with international law? Since Grotius first identified the concept that States seek esteem from the broader global community,<sup>1</sup> its parameters have proven illusive.<sup>2</sup> Nonetheless, the notion remains central to discussions of why States comply with international agreements.<sup>3</sup>

Understanding the reputational mechanism that impels State compliance is especially important to human rights treaties. Unlike other regimes, States that ratify and abide by the terms of these instruments receive neither reciprocal nor immediate benefits.<sup>4</sup>

---

\* Professor and Associate Dean, Valparaiso University School of Law; Professor, Drexel University College of Law.

\*\* Cabell Research Professor, William and Mary School of Law. This Article was written while Visiting Professor, Harvard Law School (2005-06). We thank Kenneth Dau-Schmidt, Andrew Guzman, Eric Posner, Jeremy Telman, Stepan Wood and participants at a faculty workshop at the University of Connecticut School of Law for their comments; Yvonne Chan and Laura Colleton for their research assistance; and Penelope Stein and Jane Scarpellino for their support.

1. See generally HUGO GROTIUS, *THE RIGHTS OF WAR AND PEACE* (A. C. Campbell, trans., reprint ed., M. Walter Dunne 1979) (1625); see also ABRAM CHAYES & ANOTINA HANDLER CHAYES, *THE NEW SOVEREIGNTY: COMPLIANCE WITH INTERNATIONAL REGULATORY AGREEMENTS* 27-28 (1995) (arguing that compliance derives from the need to maintain one's status within a highly interrelated community of States).

2. See, e.g., Harold Hongju Koh, *Internalization Through Socialization*, 54 DUKE L.J. 975, 977 (2005) (referencing Ryan Goodman and Derek Jinks as stating that “the mechanisms of social influence have so far been grossly underspecified in the literature”) [hereinafter Koh, *Internalization*].

3. See, e.g., Harold Hongju Koh, *Why Do Nations Obey International Law?*, 106 YALE L.J. 2599, 2646 (1997) [hereinafter Koh, *Why Do Nations*] (explaining that international norm internalization occurs when a transnational actor provokes an interaction with another actor, forcing an interpretation of the applicable global norm in order to coerce the other party to internalize the new interpretation of the norm); see also George W. Downs & Michael A. Jones, *Reputation, Compliance, and International Law*, 31 J. LEGAL STUD. 95, 95 (2002) (“International relations theorists and international lawyers have argued that reputational concerns help ensure that states maintain their agreements.”); Andrew T. Guzman, *A Compliance-Based Theory of International Law*, 90 CAL. L. REV. 1825, 1835-36 (2002) (arguing that nations internalize norms to participate in the transnational legal process).

4. This is because human rights treaties control state conduct towards their own citizens. See, e.g., Oona A. Hathaway, *The Cost of Commitment*, 55 STAN. L. REV. 1821, 1823 (2003) (stating that parties to human rights treaties “receive only promises from other nations to refrain from harming their own citizens”). One could argue that positive network externalities, even to the extent of financial advantage, can be generated by human rights treaty collaboration. Such an assertion would be in line with our theory that improved reputation effects motivate international cooperation. That impact, however, is not immediate since it requires time to accrete. To the extent that monetary advantage is immediate, for example one state reworking its trade relationship with another in return for the latter ratifying a given treaty, that influence

Consequently, the desire for international esteem is the crucible by which compliance with human rights norms is determined. Professor Oona Hathaway has recently raised concerns about the efficacy of the reputational mechanism through an empirical study suggesting that ratification of human rights treaties may lead to increased human rights violations.<sup>5</sup> While these findings have been challenged on both econometric and normative grounds,<sup>6</sup> the need to understand whether Hathaway's findings are simply the result of faulty measurement, improper regime design, or the inability of social forces to compel compliance is crucial to the development and success of the international human rights project. Answers to these questions begin with the development of more detailed models of the way social forces work. Not only will such models enable us to determine if social forces can be harnessed to better ensure compliance with human rights treaties, they also provide a framework that can help us better design human rights instruments.

A number of scholars have begun to explore the social mechanism. Professors Ryan Goodman and Derek Jinks, for example, suggest that the social force is one of acculturation—a process whereby actors adopt the beliefs of the surrounding culture—and that this force stands apart from traditional notions of coercion and

---

is coercive. A more complete account of the difference between persuasion, reputation, and coercion is set forth *infra* Part III.B.

5. Oona A. Hathaway, *Between Power and Principle: An Integrated Theory of International Law*, 72 U. CHI. L. REV. 469, 526 (2005) [hereinafter Hathaway, *Between Power and Principle*]; Oona A. Hathaway, *Do Human Rights Treaties Make a Difference?*, 111 YALE L.J. 1935, 1989 (2002) [hereinafter Hathaway, *Do Human Rights Treaties*].

6. Ryan Goodman & Derek Jinks, *Measuring the Effects of Human Rights Treaties*, 14 EUROPEAN J. INT'L L. 171, 173 (2003) [hereinafter Goodman & Jinks, *Measuring the Effects*]. The essence of their critique is one of timing: human rights obligations inure on State signature and incorporation rather than on ratification. Hence, examining the period preceding implementing domestic legislation—when States are likely to be in technical violation when measured against ratification but not against incorporation—may skew the results. They also assert that Hathaway does not properly account for reservations under which States can delay implementing treaty obligations. For an empirical examination of how this type of timing differential affects perspectives regarding human rights compliance, see Linda Camp Keith, *The United Nations International Covenant on Civil and Political Rights: Does it Make a Difference in Human Rights Behavior?*, 36 J. PEACE RES. 95, 105 (1999) (“A separate analysis was conducted in which the states that derogated from the treaty were moved into the group of non-party states . . . . When this adjustment is made, the difference between states parties and non-party states personal integrity abuse increases substantially.”). See also Ryan Goodman & Derek Jinks, *How to Influence States: Socialization and International Human Rights Law*, 54 DUKE L.J. 621, 700 (2004) [hereinafter Goodman & Jinks, *How to Influence States*] (advocating the development of an empirical approach that would integrate the mechanisms of coercion, persuasion and acculturation to build an effective human rights regime); Ryan Goodman & Derek Jinks, *Toward an Institutional Theory of Sovereignty*, 55 STAN. L. REV. 1749, 1785 (2003) (arguing that universal ratification should be the first goal of treaty regimes, while achieving enforcement should come later).

persuasion.<sup>7</sup> Other commentators aver that States ratify or comply as a way of signaling their cooperative nature to others.<sup>8</sup> Dean Harold Koh suggests that a complex interaction of domestic and international processes leads to the internalization of international norms by a State.<sup>9</sup> Another school of thought, championed by Dean Anne Marie Slaughter, suggests that norms are more likely to be enforced by liberal States than by States that do not respect the rule of law.<sup>10</sup> By contrast, Professors Jack Goldsmith and Eric Posner are skeptical of social influences. They maintain that international norms are, at best, artifacts of the cooperation of States with stable and preexisting preferences.<sup>11</sup>

Each of these observations provides an interesting and valuable lens through which to view the long-recognized but misunderstood process of how desire for international esteem influences compliance. However, none provides a comprehensive framework from which specific conclusions can be drawn. Goodman

---

7. Goodman & Jinks, *How to Influence States*, *supra* note 6, at 626-27; Goodman & Jinks, *Measuring the Effects*, *supra* note 6, at 183 (suggesting that a sociological empiricism might be more helpful than Hathaway's economic approach); see also Ryan Goodman & Derek Jinks, *International Law and State Socialization: Conceptual, Empirical, and Normative Challenges*, 54 DUKE L.J. 983, 991 (2004) [hereinafter Goodman & Jinks, *International Law*] (asserting that states are influenced by acculturation); see also Jose E. Alvarez, *Do States Socialize?*, 54 DUKE L.J. 961, 962 (2004) (arguing that Brian Simpson's work on the European Convention of Human Rights would provide a valuable case study to support Goodman & Jinks' acculturation theory); Koh, *Internalization*, *supra* note 2, at 980 (suggesting that Goodman & Jinks' acculturation theory is an "intermediate way between persuasion and coercion").

8. See, e.g., Daniel A. Farber, *Rights as Signals*, 31 J. LEGAL STUD. 83, 88 (2002) (quoting Beth Simmons as arguing that governments comply with their legal commitments in order to preserve their reputations); David H. Moore, *A Signaling Theory of Human Rights Compliance*, 97 NW. U. L. REV. 879, 904 (2003) (suggesting that failing to ratify a treaty indicates antagonism towards other participants); Charles K. Whitehead, *What's Your Sign?—International Norms, Signals, and Compliance*, 27 Mich. J. Int'l L. 695 (2006). See also, Andrew T. Guzman, *The Promise of International Law*, 92 VA. L. REV. 533, 549 (2006) (asserting that states that comply with legal rules develop reputations as cooperative actors).

9. "[T]ransnational legal process is normative, dynamic, and constitutive... future transactions will further internalize those norms; and eventually, repeated participation in the process will help to reconstitute the interests and even the identities of the participants in the process." Koh, *Why Do Nations*, *supra* note 3, at 2646; see also Koh, *Internalization*, *supra* note 2, at 981 (asserting that a State's norm-internalization could be tracked "from coercion to persuasion to internal acceptance"); Harold Hongju Koh, *Transnational Legal Process*, 75 NEB. L. REV. 181, 204 (1996) (arguing that the combination of interaction between States and internalization of norms through domestic decision-making is what causes international law to stick).

10. ANNE-MARIE SLAUGHTER, *A NEW WORLD ORDER* 15 (2004) (arguing that the world order is composed of "an intricate three-dimensional web of links between disaggregated state institutions").

11. JACK L. GOLDSMITH & ERIC A. POSNER, *THE LIMITS OF INTERNATIONAL LAW* 9-10, 13 (2005); Jack Goldsmith & Eric A. Posner, *Moral and Legal Rhetoric in International Relations: A Rational Choice Perspective*, 31 J. LEGAL STUD. 115, 119 (2002) (arguing that States use international legal and moral rhetoric while pursuing their own interests).

and Jinks, for instance, identify a number of different processes by which acculturation occurs but cannot identify what process will affect State behavior in any particular situation.<sup>12</sup> In arguing that internalization is the key to compliance, Koh also provides little more than descriptive guidance of how such a process works, relying instead on the above-stated claim that internalization results from a complex and dynamic interaction of States.<sup>13</sup> Signaling theory similarly provides scant direction for effective regime design and implementation.<sup>14</sup> Likewise, liberal theorists cannot provide a full understanding of why democratic States are more compliant than others.<sup>15</sup> The result is an important body of work that recognizes the centrality and implications of esteem-based mechanisms on compliance, but fails to provide a level of detail necessary to understand or predict how the desire for esteem affects State behavior.

This Article provides a more complete understanding of the forces behind treaty creation and compliance by drawing on the developing domestic expressive law literature. By expressive law, we mean the impact that law and legal process have on individual behavior (construed heuristically as personal or State behavior) by affecting the social, or normative, meaning of that behavior.<sup>16</sup> We

---

12. See, e.g., Goodman & Jinks, *International Law*, *supra* note 7, at 984 (agreeing with criticisms of their theory of acculturation by Koh and Alvarez that “more fine-grained empirical work is required to document the specific causal pathways by which law influences States”).

13. Eric A. Posner, *International Law and the Disaggregated State*, 32 FLA. ST. U. L. REV. 797, 800-01 (2005) (arguing that constructivist scholars such as Harold Koh and Anne-Marie Slaughter rely heavily on the notion of internalization but have not yet developed a coherent theory of how internalization happens).

14. In particular, existing descriptions of signaling fail to describe how nations agree on which behaviors become signals. For a fuller discussion, see *infra* Part III.D.

15. See, e.g., Peter M. Haas, *Why Comply, or Some Hypotheses in Search of an Analyst*, in INTERNATIONAL COMPLIANCE WITH NONBINDING ACCORDS 21, 36-39 (Edith Brown Weiss, ed., 1997) (positing an intuitionist theory that wealthier liberal states are more likely to comply with international law than their poorer peers without explaining why this is so); Daniel E. Ho, *Compliance and International Soft Law: Why Do Countries Implement the Basle Accord?*, 5 J. INT’L ECON. L. 647, 682 (2002) (challenging liberal theorists on normative grounds, while also demonstrating the absence of empirically verifiable evidence); Anne-Marie Slaughter, *International Law in a World of Liberal States*, 6 EUR. J. INT’L L. 503, 535 (1995) (theorizing that aggregation of preferences among influential domestic actors influences liberal democracies to comply to a greater extent than others).

16. Throughout the Article we follow the prevailing convention of referring to States as individuals when analyzing actions that usually result from aggregate decision-making. For a fuller account, see Goodman & Jinks, *How to Influence States*, *supra* note 6, at 646 (“There is good reason to question whether states” are influenced by social forces. “After all, much of the research suggesting the presence of this mechanism centers on the cognitive processes of individuals. . . . Substantial evidence strongly suggests, on both counts, that they do.”); see also William Bradford, *International Legal Compliance: An Annotated Bibliography 4* (unpublished manuscript, on file with authors) (identifying a body of scholarship that regards the State as an



model the effects of international lawmaking on the beliefs of States and ultimately analyze how changing normative belief can influence the expected utility of decisions to ratify and/or comply with international law. Along the way, we also demonstrate how our model clarifies a number of debated issues within the developing compliance literature.<sup>17</sup>

We draw a number of conclusions relevant to the structure of treaty regimes from our analysis and provide a model of how expressive influences can best be harnessed to serve the goals of ratification and compliance. In general, commentators have considered the influences motivating compliance too narrowly by concentrating on the relationships surrounding the decision of whether to ratify or obey a given individual treaty. Rather than being atomistic, the pull of international society is animated by a need-reinforcement notion that is contingent on a larger system of global cooperation. Absent these broader socializing mechanisms, the ability to harness norms to achieve treaty compliance will be minimal. Finally, we describe how treaty formation may affect the willingness of rebel States to act in contravention of international norms, thus lessening the ability of a treaty regime to achieve its ultimate goals.

The Article proceeds as follows: Part I sets forth our expressive theory of international law. Beginning from a standard law and economics baseline that assumes individuals act rationally to maximize their own utility, we apply expressive law principles to explain how normative pressure influences rational actors to alter their behavior and beliefs while seeking esteem from other group members. We call this the expressive model. Part II applies this framework to State international treaty ratification and compliance. It describes how international law may affect States' estimates of the likelihood of being sanctioned by other States and may also affect preferences regarding the behaviors being regulated. Part II also considers how international law may affect domestic lawmaking, as well as the market behavior of citizens in a country that ratifies a particular treaty. Ultimately, we describe how normative forces may indeed produce greater compliance in liberal States. Part III returns

---

abstraction without the capacity to exercise choice between alternatives and thus treats questions of compliance as one of human agency). Because this convention does not adequately account for the varied roles that individuals play in reaching group-based decisions, for instance the weight given by a State to the opinions of experts or epistemic communities, we will return to these more nuanced effects where appropriate.

17. For our previous, respective, ventures into expressive law scholarship, see Alex Geisinger, *A Belief Change Theory of Expressive Law*, 88 IOWA L. REV. 35 (2002); Michael Asbley Stein, *Under the Empirical Radar: An Initial Expressive Law Analysis of the ADA*, 90 VA. L. REV. 1151 (2004).

to the international law literature and demonstrates how an expressive theory harmonizes the various schools of thought into a more comprehensive view of international law ratification and compliance. The Article concludes by providing an example of the way in which enforcing our theory would affect monitoring devices in an evolving United Nations human rights treaty.

## I. EXPRESSIVE THEORY

Expressive law examines law's potential for changing the social meaning of particular behavior by altering the social cost of undertaking that behavior. The decision to act is based on a rational assessment of the consequences of the behavior and the certainty of belief.<sup>18</sup> In the context of international law and process, the need-reinforcement principle adds to the basic rational choice model of behavior by providing a basis for understanding and modeling how the desire for esteem influences a State's assessment of the utility of entering into and abiding by any particular legal regime.

### A. *The Reasoned-Action Model*

Traditional law and economics literature operates from the premise that people act rationally to maximize their own utility when choosing among alternatively available courses of conduct.<sup>19</sup> Under this framework, known as rational choice theory, law operates by varying the cost to an individual of satisfying her preferences through the use of exogenous sanctions.<sup>20</sup> For any given opportunity, increasing the associated cost will decrease that individual's desire to choose that opportunity; conversely, a decrease in cost will encourage

---

18. See generally Robert Cooter, *Expressive Law and Economics*, 27 J. LEGAL STUD. 585, 586 (1998) (arguing that rational people will internalize a norm if that action will create a benefit for them); Lawrence Lessig, *The Regulation of Social Meaning*, 62 U. CHI. L. REV. 943, 1001 (1995) [hereinafter, Lessig, *The Regulation*] (discussing how “[s]ocial meanings are part of the benefits and costs associated with any individual action”); Richard H. McAdams, *A Focal Point Theory of Expressive Law*, 86 VA. L. REV. 1649, 1651 (2000) (explaining that when individuals have an interest in coordination, law can guide behavior by “influencing expectations” of how others will act); Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2025 (1996) (advocating that “an expressive approach to law depends on an assessment of social consequences”).

19. See WILLIAM J. BAUMOL & ALAN S. BLINDER, *ECONOMICS: PRINCIPLES AND POLICY* 99-103 (8th ed. 2000) (explaining that a person will choose to make a purchase based on the utility the person will gain from the purchase); see also MARK PERLMAN & CHARLES R. MCCANN JR., *THE PILLARS OF ECONOMIC UNDERSTANDING: FACTORS AND MARKETS* 301 (2000) (describing the market as a “model of allocative efficiency”).

20. “In the vision of law that dominates economics-influenced legal theory, law imposes sanctions to solve problems.” McAdams, *supra* note 18, at 1650. McAdams uses this axiom as a departure point for his version of expressive law theory.

the individual to satisfy her desire by choosing that opportunity.<sup>21</sup> In other words, manipulating the opportunity set available to a given actor will alter her subsequent choices. This standard economic account has proved a useful baseline method for modeling human behavior and thereby predicting the effects of particular policies.<sup>22</sup>

Expressive law seeks to understand law's potential for changing the social meaning of particular behavior and thereby to alter the social cost of undertaking that behavior.<sup>23</sup> Some authors consider the social, or symbolic, meaning of certain legal doctrines or decisions; others consider the impact that law may have on mediating the social meaning of an activity.<sup>24</sup> What is crucial to this analysis is the nexus between law, norms, and social meaning. When designed appropriately, law can cause individuals to alter their own behavior because it either induces them to change their tastes (internalization), or it creates a fear of bearing social sanctions (second order sanctions).<sup>25</sup>

Our expressive theory of international law is based on the Reasoned-Action model of decision-making, which identifies two factors that affect an individual's decision to undertake a behavior.<sup>26</sup> The first factor is the individual's attitude toward the behavior itself. The second factor is the individual's beliefs about what other people think of the behavior. The Reasoned-Action model is diagrammed in Figure 1.<sup>27</sup>

---

21. For a discussion of this point made in the context of criminal activity, see Kenneth G. Dau-Schmidt, *An Economic Analysis of the Criminal Law as a Preference-Shaping Policy*, 1990 DUKE L.J. 1, 2-3, 5.

22. Two Nobel laureates defend this model from the perspective that preferences are relatively static and that studying variable taste is a futile endeavor. See George J. Stigler & Gary S. Becker, *De Gustibus Non Est Disputandum*, 67 AM. ECON. REV. 76, 76 (1977).

23. See, e.g., Cooter, *supra* note 18, at 585-88 (describing the role of law in the development of social norms, and socioeconomic law and economics, which seeks to inject psychological and social factors related to wealth and race into otherwise "neutral" economic analyses); Lessig, *The Regulation*, *supra* note 18, at 946-48 (examining the social construction of orthodoxy and its place in the law); McAdams, *supra* note 18, at 1650-51 (suggesting that law may be alternatively conceptualized for its expressive functions, as well as its traditionally acknowledged enforcement functions); Sunstein, *supra* note 18, at 2024-25 (considering how legal statements might be designed to change social norms).

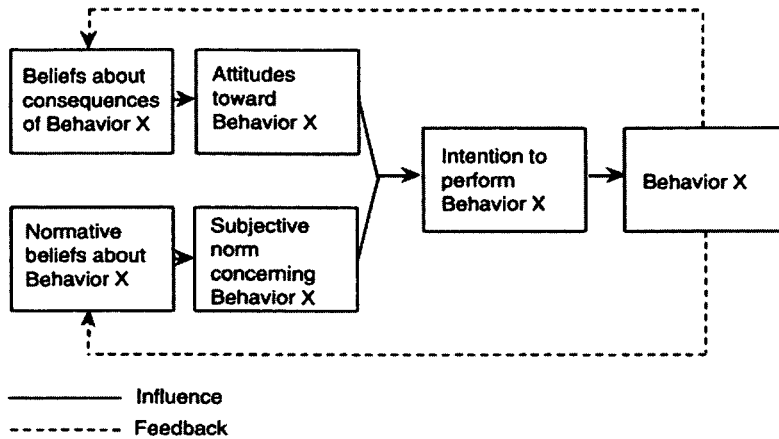
24. Lawrence Lessig, *The New Chicago School*, 27 J. LEGAL STUD. 661, 681-83 (1998).

25. The above account is synthesized and abbreviated. For a review of the development of the field and the nuances contained therein, see Geisinger, *supra* note 17, at 44-55.

26. MARTIN FISHBEIN & ICEK AJZEN, BELIEF, ATTITUDE, INTENTION AND BEHAVIOR: AN INTRODUCTION TO THEORY AND RESEARCH 13-18 (1975).

27. For additional sources of support for the reasoned action model, see Icek Ajzen & Martin Fishbein, *Attitudinal and Normative Variables as Predictors of Specific Behavior*, 27 J. PERSONALITY AND SOC. PSYCHOL. 41, 42 (1973) (describing a theory that predicts behavior using both an "attitudinal" and a "normative" factor); Icek Ajzen & Martin Fishbein, *The Prediction of Behavior from Attitudinal and Normative Variables*, 6 J. EXPERIMENTAL SOC. PSYCHOL. 466, 466-67 (1970) (reporting the results of an experiment analyzing how one's intent to act can be

Figure 1



As the model suggests, an individual decides whether to engage in particular behaviors by reasoning about how (a) good or bad, and (b) likely or unlikely, the outcomes associated with a given behavior will be (the “behavioral attitude”). An individual will also consider the amount and quality of social pressure to engage or not engage in that specific behavior (the “subjective norm”). The behavioral attitude and subjective norm together determine the individual’s intent to act.<sup>28</sup> Thus, understanding one’s attitude toward a behavior and one’s belief

---

predicted by one’s attitude and normative beliefs). For a recent overview of research in this area, see Icek Ajzen, *The Nature and Operation of Attitudes*, 52 ANN. REV. OF PSYCHOL. 27, 28 (2001) (reviewing research published between 1996 and 1999). Figure 1 is derived from RUSSELL VEITCH & DANIEL ARKKELIN, ENVIRONMENTAL PSYCHOLOGY: AN INTERDISCIPLINARY PERSPECTIVE 109 (1995).

28. See VEITCH & ARKKELIN, *supra* note 27, at 108-11 (explaining the theory and relating it to environmental perception). One interesting aspect of the model is that it helps us understand when attitude and behavior are inconsistent; that is, when one is predisposed positively toward a behavior but still does not undertake the behavior due to subjective expectations regarding social pressure.

about the subjective norm can help to determine<sup>29</sup> one's desire to undertake that behavior.<sup>30</sup>

### *B. Beliefs as the Building Blocks of Attitude*

While the subjective norm is defined in terms of one's beliefs about what others think of a behavior, it is more difficult to conceive of the concept of attitude in terms of an individual's beliefs.<sup>31</sup> Attitude toward a behavior can, however, be defined as a function of what individuals believe about the consequences of the behavior,<sup>32</sup> the certainty of their beliefs, and their evaluations (either positive or negative)<sup>33</sup> of those consequences.<sup>34</sup> This relationship can be expressed

---

29. The model itself can be deceptively simple. In particular, the model conceives of the individual in a vacuum, uninfluenced by social context. Intentions to act, of course, rely significantly on social context. For example, an individual may have different attitudes toward an activity based on the normative group to which she belongs. Criminals, for example, think differently about crime than police and an individual may have a different attitude toward pollution as a member of her business community than she does as a member of her home or family community. For a general analysis of the influence of social context on norms, see ATTITUDES, BEHAVIOR, AND SOCIAL CONTEXT: THE ROLE OF NORMS AND GROUP MEMBERSHIP 2 (Deborah J. Terry & Michael A. Hogg eds., 2000).

30. Note that intending to undertake a behavior and actually acting are not always the same. There may be physical limitations to behavior. Thus, I may desire to climb a mountain but weather, geography or physical exhaustion may keep me from so doing. See generally Lessig, *The Regulation*, *supra* note 18, at 955-57 (noting that physical limitations may keep us from doing what we want).

31. See VEITCH & ARKKELIN, *supra* note 27, at 109. We will discuss in much greater detail the subjective norm *infra* Parts II.B.1-2.

32. One potentially significant limitation on the effective use of a belief-based theory is the fact that any behavior is associated with a virtually limitless number of beliefs, thus significantly limiting the ability to analyze the effect law will have on attitude. However, only a relatively small number of beliefs affect our attitude. Due to limited attention span, apprehension, and information processing abilities, individuals can only process a small number of beliefs at any single time. Thus, although an individual may have a large number of beliefs that, if given time, she could recall about a particular behavior and its consequences, only a maximum of between five and nine of these beliefs underlie her attitude. FISHBEIN & AJZEN, *supra* note 26, at 218.

33. Evaluation of a consequence means simply that one thinks positively or negatively about the consequence of an action. Consider one's attitude toward wearing a seatbelt. Certain consequences are generally positively evaluated (e.g. safety), while certain other consequences are generally negatively evaluated (e.g. discomfort). Evaluations of consequences are formed by standard processes of conditioning. See, e.g., *id.* at 277 (noting that evaluations, in the end, must be accounted for by the process of conditioning). These processes include operant conditioning, classical conditioning, and vicarious conditioning. See VEITCH & ARKKELIN, *supra* note 27, at 105-07 (defining and providing examples of types of conditioning).

34. The elements of the belief-based theory are as follows:

- (1) An individual holds many beliefs about a given object; i.e., the object may be seen as related to various attributes, such as other objects, characteristics, goals, etc.
- (2) Associated with each of the attributes is an implicit evaluative response, i.e., an attitude.
- (3) Through conditioning, the evaluative responses are associated with the attitude object.
- (4) The conditioned evaluative

by the equation  $A_o = \sum b_i e_i$  where A is the attitude toward behavior O; b is the belief about O; (i.e. the subjective certainty that O will result in consequence i); e is the evaluation of the consequence; and n is the number of beliefs.

This theory of beliefs as the basis of attitude can be correlated with the Subjective Expected Utility theory of behavioral science. According to this concept, “[W]hen a person has to make a behavioral choice, he will select that alternative which has the highest subjective expected utility,” i.e., the alternative which is likely to lead to the most favorable outcomes. This can be stated as  $SEU = \sum SP_i U_i$  where “ $SP_i$  is the subjective probability that the choice of this alternative will lead to some outcome i, and  $U_i$  is the utility of the outcome i.” This model can be recast in terms of beliefs about consequences. That is,  $SP = b$  and  $U = e$ , or  $A_o = \sum b_i e_i$ .<sup>35</sup>

Take, as an example, a State’s attitude toward administrative detention, a procedure whereby individuals are held without being charged or tried with criminal violations.<sup>36</sup> States may evaluate the

responses summate, and thus (5) on future occasions the attitude object will elicit this summated evaluative response, i.e., the overall attitude.

FISHBEIN & AJZEN, *supra* note 26, at 29. The theory of belief-based attitude and intent has its roots in the earliest work of Professor Fishbein. See, e.g., Martin Fishbein, *An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward That Object*, 16 HUM. REL. 233, 234 (1963) (positing that all beliefs about an object are related to an individual’s attitude, since all beliefs about an object contain an evaluative aspect). For a description of the belief-based theory of attitude and intent formation, this Article will rely primarily on FISHBEIN & AJZEN, *supra* note 26, which remains the most comprehensive exegesis of the theory. It should, however, be noted that the theory has been elaborated further in a number of articles—sometimes responding to criticism—by Professors Fishbein, Ajzen, and others. For examples of criticism, see Vernon E. Cronen & Richard L. Conville, *Fishbein’s Conception of Belief Strength: A Theoretical, Methodological and Experimental Critique*, 42(2) SPEECH MONOGRAPHS 143, 147-150 (1975) (presenting three studies that cast doubt on Fishbein’s theory of attitude); Joseph R. Priester & Monique A. Fleming, *Artifact or Meaningful Theoretical Constructs?: Examining Evidence for Nonbelief- and Belief-Based Attitude Change Processes*, 6(1) J. CONSUMER PSYCHOL. 67, 69 (1997) (noting and describing the evidence for a number of nonbelief-based attitude change processes). For further exposition and response to these criticisms, see Martin Fishbein & Susan Middlestadt, *Noncognitive Effects on Attitude Formation and Change: Fact or Artifact?*, 4(2) J. OF CONSUMER PSYCHOL. 181, 184 (1995) (analyzing a number of critical studies and arguing that the contribution of factors, other than belief-based expectancy-value measures to the prediction of attitude, can be seen as a methodological artifact of using inappropriate measures); see also Martin Fishbein & Susan E. Middlestadt, *A Striking Lack of Evidence for Nonbelief-Based Attitude Formation and Change: A Response to Five Commentaries*, 6(1) J. CONSUMER PSYCHOL. 107, 108 (1997) (arguing that most criticism avoids assessing the belief-based structure that underlies attitude formation).

35. FISHBEIN & AJZEN, *supra* note 26, at 30-31; see also Lynn R. Anderson & Martin Fishbein, *Prediction of Attitude From the Number, Strength, and Evaluative Aspect of Beliefs About the Attitude Object: A Comparison of Summation and Congruity Theories*, 2(3) J. PERSONALITY & SOC. PSYCHOL. 437 (1965) (arguing that basic summation of belief and evaluation yields significantly better predictions of attitude than congruity theory).

36. For practical examples and further analysis, see Amnesty International’s website, <http://web.amnesty.org/pages/isr-action-detention> (last visited Nov. 4, 2006).

procedure both positively as providing greater protection against terrorist activities and negatively as reducing civil liberties. The certainty with which a State holds these beliefs, in conjunction with its evaluations of the effects of each of these outcomes, can determine that country's attitude regarding the behavior. To see why, assume a scale of certainty that runs from 0 (no certainty) to +100 (strong certainty) and a similar scale for evaluation -100 (strong dislike) to +100 (strong like). Applying these factors to administrative detention could have the results set forth in Figure 2.

Figure 2

Belief	Certainty	Evaluation	be
Protects against Terrorist Acts	+50	+50	+2500
Restricts Civil Liberties	+40	-40	-1600
		$A_o = \sum b_i e_i$	= +900

Based on these beliefs, a State would be inclined to use administrative detention, but not very strongly (because the utility is only +900).

### C. The Need-Reinforcement Principle

The notion that States value their reputation among other States is, of course, not new to discussions of international law.<sup>37</sup> Early proponents of this view discussed the "need to preserve solidarity with one's fellow States as an explanation for compliance."<sup>38</sup> These scholars discussed the incentive to comply with international law as based on prodding from other States with whom that State is

---

37. See, e.g., Downs & Jones, *supra* note 3, at 96 (arguing that states have multiple reputations which vary in connection with different agreements, and thus, while reputation does matter, the standard notion that states keep commitments out of concern for their reputations may overestimate its impact); Hathaway, *Between Power and Principle*, *supra* note 5, at 494 (discussing reputation as a consideration that influences States' decisions whether to commit to and comply with treaties); Colin B. Picker, *Reputational Fallacies in International Law: A Comparative Review of United States and Canadian Trade Actions*, 30 BROOK. J. INT'L L. 67, 69-70 (2004) ("The potential harm of reputation derives from the fact that state reputation is an important and vital component in the smooth operation of international law. With few traditional legal mechanisms to ensure that states comply with international law, it is arguable that reputation is an important, even crucial, mechanism for securing state compliance.").

38. Koh, *Why Do Nations*, *supra* note 3, at 2613 (internal citations omitted); see also Nathaniel Berman, *The Paradoxes of Legitimacy: Case Studies in International Legal Modernism*, 32 HARV. INT'L L.J. 583, 585 (1991) (describing Alfred Verdross's theory that obligation in international law must be grounded in a basic norm, which is "the fundamental rule whose function is to institute the supreme creative authority of law").

engaged.<sup>39</sup> Others still suggested that States shared a “common consciousness” that led them to comply with international law because “particular rules are nested within a much broader fabric of ongoing communal relations.”<sup>40</sup>

More recently, international relations theorists have focused on interdependence as the underlying mechanism through which reputational factors affect state action.<sup>41</sup> These theorists view international law as reflecting the strategies engaged in by States involved in long-term relationships to increase their esteem within the global community.<sup>42</sup> Social norms and the rules that embody them, they assume, influence the strategies pursued by States within these games.<sup>43</sup> Our expressive theory builds on and bridges some of these understandings. It provides a definite understanding of the forces that lead to the pull of international society that many scholars have identified as a key to compliance. Further, it bases the desire for reputational benefits in a preference for esteem that develops from beneficial interaction with other members of a group.

Most law and economics scholars conceive of norms as arising from cooperation problems that confront rational individuals acting in their own self-interest.<sup>44</sup> A classic example of a cooperation problem is

---

39. See Koh, *Why Do Nations*, *supra* note 3, at 2613 (“[T]he interwar years modified the process-based strand of thinking about the compliance question by mixing *process with reputation*: the ‘solidaristic’ strand that emerged derived a nation’s incentive to obey from the encouragement and prodding of other nations with whom it is engaged in a managerial, discursive legal process.” (emphasis in original)).

40. *Id.* at 2617 (arguing that English scholars developed the notion of a “common consciousness” in invoking the Grotian notion of “international society,” within which nations comply for communitarian reasons).

41. SLAUGHTER, *supra* note 10, at 170 (noting that government networks “could develop and support the implementation of ‘network norms’ that would strengthen the integrity and competence of all their members”). As explained by Koh, “scholars began to look as well at transnational networks among nonstate actors, international institutions, and domestic political structures as important mediating forces in international society . . . . They understood compliance with international law to result almost entirely from the functional benefits such compliance provides.” Koh, *Why Do Nations*, *supra* note 3, at 2624-25.

42. See, e.g., Downs & Jones, *supra* note 3, at S96 (averring that the reputational consequences of defection require States to engage in a complex cost-benefit analysis when deciding whether to comply with their obligations); Anne E. Sartori, *The Might of the Pen: A Reputational Theory of Communication in International Disputes*, 56 INT’L ORG. 121, 123 (2002) (noting that States behave with a view to how the reputations they engender will impact future and repeated transactions).

43. See generally Mark A. Chinen, *Game Theory and Customary International Law: A Response to Professors Goldsmith and Posner*, 23 MICH. J. INT’L L. 143, 167 (2001) (extrapolating game theory to international legal regimes).

44. See, e.g., Thomas F. Cotter, *Legal Pragmatism and the Law and Economics Movement*, 84 GEO. L.J. 2071, 2126 n.235 (1996) (noting that individuals may be hesitant to act in an altruistic fashion unless there is some assurance that others will be compelled to act in a similar fashion); Steven A. Hetcher, *Creating Safe Social Norms in a Dangerous World*, 73 S. CAL. L. REV. 1, 7 n.21 (1999) (quoting Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L.



the “prisoner’s dilemma,” which posits two rational, self-interested individuals who must choose between alternate strategies. In the circumstance of the game, pursuit of individual self interest leads to worse results for each individual than if he or she had cooperated with the other.<sup>45</sup>

Take, for example, the following scenario between prisoners Row and Column, who have been placed in separate cells at the police station and are being questioned. If one inmate tells on the other, the tattler will be let off for cooperation, and the other detainee will get a three-year sentence. If neither tells, each will be found guilty of a lesser offense and serve one year in prison. If both tell, each will be convicted of a more significant offense and incarcerated for two years. The options and consequences can be diagrammed as follows.

Figure 3

	Cooperate (withhold)	Defect (tell)
Cooperate (withhold)	1/1	3/0
Defect (tell)	0/3	2/2

Given these circumstances, Row will always tell. Here’s why: Assume first that Column does tell. If Row fails to do likewise, he will get three years in jail; if he does tell, he will receive only a two-year sentence. Now assume that Column does not tell. Row will not be imprisoned if he tells, but he will be punished with one year in jail if

---

REV. 903, 918 (1996) (“Good social norms solve collective action problems by encouraging people to do useful things that they would not do without the relevant norms.”); Steven A. Hetcher, *Norm Proselytizers Create a Privacy Entitlement in Cyberspace*, 16 BERKELEY TECH. L.J. 877, 902-03 n.90 (2001) (describing four alternative accounts of how norms may serve as solutions to iterated collective action problems); Eric A. Posner, *Symbols, Signals, and Social Norms in Politics and the Law*, 27 J. LEGAL STUD. 765, 797 (1998) (“[T]he social norm describes the behavior that arises in equilibrium. It is not that X punishes Y for violating a social norm; rather, X (and many other people) avoids Y because Y’s behavior reveals to X that association with Y will not serve X’s interests.”); Elmer J. Schaefer, *Predicting Defection*, 36 U. RICH. L. REV. 443, 462 (2002) (discussing social norms, specifically signaling, as an answer to defections in prisoner’s dilemma problems with cooperation theories).

45. The most successful strategy is the well-known “tit-for-tat” strategy that emerged victorious from a number of computer tournaments run by Robert Axelrod. Following the tournaments, Axelrod simulated natural selection with sixty-three programs by adjusting the number of offspring produced in each successive round based on a strategy’s performance in the previous round. After one thousand rounds of play, weak programs became extinct, and so did some predatory programs that had survived by exploiting dwindling programs lower in the food chain. Interestingly, in this game, which was designed to simulate Darwinian natural selection, TFT won again, just as it had in Axelrod’s tournaments. See generally David Crump, *Game Theory, Legislation, and the Multiple Meanings of Equality*, 38 HARV. J. ON LEGIS. 331, 377-78 (2001).

he remains silent. In these circumstances, it is better for the self-interested Row to tell regardless of what Column does. The dominant strategy for both players will therefore be to tell, resulting in each getting locked up for two years. By contrast, if neither tells, each gets only one year in jail. Consequently, the pursuit of individual self-interest by Row and Column leads to worse results than if they had cooperated and both withheld information.

While defection is the dominant strategy in a one-time play of the prisoner's dilemma, cooperation is a natural result of such a problem in situations where the parties will play the game a substantial number of times.<sup>46</sup> Let us assume that Column and Row are now, respectively, a wholesaler and retailer of goods. They want to create a relationship whereby Column will supply the goods at a certain cost. If Column delivers the quality of goods agreed upon, both parties will make two. If Column cheats and sends goods of lesser quality he will make three and Row will make zero, but Row will end their relationship and Column will have to look for other cooperative partners. A similar result would occur if Row cheats, for example, by challenging the quality of the goods and withholding full payment. Assuming a desire to play for a number of times, it is better for the parties to cooperate than to sever their ties. This is because making two regularly is better than making three a few times while also developing a reputation for being untrustworthy and so losing future cooperative opportunities. As Professor Eric Posner has pointed out, "logic shows that the optimal move is always to cooperate."<sup>47</sup> Norms are, in turn, artifacts of the long-term cooperation of these rational individuals.

The game-theoretic model of norm formation is, of course, extremely parsimonious. In particular, it does little to help us identify specifically the way in which individual beliefs or preferences can be linked to the behavioral standards embodied by norms. Nor does the model provide an understanding of why individuals comply with norms. We will use the concept of need reinforcement in conjunction with the basic social psychology of norms and groups to inform the rational choice model, and in particular provide an understanding of these issues. By identifying norms as reflections of aggregate preference and normative behavior as a signal of the importance of group standing to an individual, the model provides a framework for

---

46. ERIC A. POSNER, *LAW AND SOCIAL NORMS* 19 (Harvard Univ. Press 2000) (developing the foundation for a theory of norms as signals of one's cooperativeness).

47. *Id.* at 16. Posner also suggests that the logic of cooperation extends to games involving more than two players by assuming that everyone has sufficient information about other people's past activities. *Id.* Thus, defection from one pairwise transaction will not lead to a "clean slate" in the next pairwise transaction. *Id.*

considering how estimations of group preference may inform and effect behavior.

The sociological model leads to a particular view of groups and norms. Pursuant to the rational choice perspective, groups form when individuals come together for the mutual satisfaction of their own needs.<sup>48</sup> The individual is the basic unit of such a conception of the group,<sup>49</sup> and interdependence is the force that holds these individuals together.<sup>50</sup> In this sense, the group is simply a reflection, or aggregation, of the individuals that comprise it, and the idea of a group as something other than a collection of individuals is meaningless.<sup>51</sup> The idea of a social norm within this framework is, in turn, simply the reflection of the aggregate preferences of the individuals that comprise the group.<sup>52</sup> That is, norms are the reflection of the perceived majority position of any group of individuals and can be determined by simply combining the individual positions of the majority of group members.<sup>53</sup>

It is difficult, however, to reconcile this view of normative behavioral standards with the notion that normative behavior provides information on one's willingness to cooperate with other group members. The connection between certain moral norms, such as "do unto others as they would do unto you," and one's cooperativeness, is apparent. It becomes harder, however, to see the relationship between other norms—for instance eating hot dogs at a baseball game—and one's cooperative nature.<sup>54</sup>

---

48. See, e.g., MUZAFER SHERIF, *GROUP CONFLICT AND COOPERATION: THEIR SOCIAL PSYCHOLOGY* 64-69 (1966) (illustrating how shared identity and group organization arise as derivative phenomena from interdependence between group members).

49. This concept has its roots in some of the earliest work of social psychology. As early as 1924, psychologists argued that the individual was the only psychological reality and that there was nothing in the group that was not in the individual. FLOYD HENRY ALLPORT, *SOCIAL PSYCHOLOGY* 4 (The Riverside Press 1924) ("There is no psychology of groups which is not essentially and entirely a psychology of individuals.").

50. John C. Turner, *Social Categorization and the Self-Concept: A Social Cognitive Theory of Group Behavior*, 2 *ADVANCES IN GROUP PROCESSES* 77, 79 (1985).

51. See ALLPORT, *supra* note 49, at 4 (arguing that the individual is the only psychological reality and that there is nothing in the group that is not in the individual).

52. See Turner, *supra* note 50, at 79 (discussing how groups are defined in terms of shared values that prescribe beliefs, attitudes and conduct).

53. *Id.* at 82 (observing that various theories of social influence have been characterized by individualism, as evidenced by the assumption that the social norm is assumed to be exactly the sum of its parts).

54. See Richard H. McAdams, *Signaling Discount Rates: Law, Norms and Economic Methodology*, 110 *YALE L.J.* 625, 676-78 (2001) (reviewing ERIC A. POSNER, *LAW AND SOCIAL NORMS*, *supra* note 46) [hereinafter McAdams, *Signaling Discount Rates*]. He asserts that it would be inefficient for all of these behaviors to act as signals of cooperativeness. Instead, he suggests, the most efficient way to create a reputation for cooperativeness is, simply, to cooperate with others.

Posner has attempted to solve this problem by describing norms as behavioral equilibria that result from people signaling their discount rates to one another.<sup>55</sup> He suggests that preferences regarding the value of future payoffs differ among the population. Thus, people with low discount rates are less likely to defect from a cooperation game because they value future payoffs higher than most. Posner deems such people “good types.”<sup>56</sup> In order to distinguish themselves from bad types, good types engage in behaviors that signal their higher discount rates. Because they value future payoffs more highly, good types are willing to undertake more expensive signaling behaviors. Norms, to Posner, are the behavioral equilibria that result from good and bad types signaling their discount rates.<sup>57</sup> While Posner’s effort continues to be the most comprehensive attempt to explain norm formation, and, in particular, to explain the normative basis for a number of specific behaviors, it has been subject to criticism.<sup>58</sup> We discuss below one major criticism concerning the inability of the theory to explain how States choose which behaviors do and do not act as signals.<sup>59</sup>

The social-psychological model provides a simpler explanation of how norms reflect a State’s cooperativeness. This explanation is rooted in the mutual attraction that arises between States that are interdependent. This attraction is rooted in the operation of a need satisfaction or “reinforcement” principle: mutual liking between group members reflects the extent to which positive, gratifying or rewarding outcomes are associated directly or indirectly with being in a cooperative relationship with each other.<sup>60</sup> The greater the perceived rewards of group membership, the greater the attraction to the group, and the less likelihood of defection.

---

55. POSNER, *supra* note 46, at 19-23.

56. *Id.* at 18-19 (noting that a good type is more likely to cooperate in a repeated prisoner’s dilemma than a bad type is, because the good type cares more about the future payoffs that are lost if cooperation fails).

57. *Id.* at 18-22 (finding the origin of the variation of social norms across time and place in the struggle by good types to differentiate themselves from bad types via signaling).

58. See, e.g., Steven A. Hetcher, *Commentaries on Eric Posner’s Law and Social Norms: Cyberian Signals*, 36 U. RICH. L. REV. 327, 359 (2002) (“Posner errs, then, by using signaling of discount rates as the sole explanation for norms.”); Dan M. Kahan, *Commentaries on Eric Posner’s Law and Social Norms: Signaling or Reciprocating? A Response to Eric Posner’s Law and Social Norms*, 36 U. RICH. L. REV. 367, 371 (2002) (“Posner’s signaling model, as it stands . . . should not be trusted to guide public policy even provisionally for four distinct reasons.” (emphasis in original)); see also McAdams, *Signaling Discount Rates*, *supra* note 54, at 627 (stating that Posner’s discount-rate signaling theory of social norms exaggerates the importance of the motivation and mechanisms it employs and understates the importance of the normative origins of norms).

59. See *infra* Part III.D.

60. Turner, *supra* note 50, at 79.

Normative pressure is, in turn, an external force that affects a State's behavior only to the extent it is concerned about other States to which it is attracted.<sup>61</sup> Put simply, if a State wants to do something it perceives is not condoned by other group members, and there is a sense of mutual liking or attraction between that State and the other group members, the State risks disapproval from others to which it is attracted.<sup>62</sup> A group member who seeks esteem is thus required to estimate which behaviors are approved by other group members. The more uniformly held and highly valued the preference, the more likely it will assert normative force.

Let us return to the example of administrative detention and a given State's belief about group preferences regarding this behavior. Consider two different possible levels of belief regarding the norm. In one case, imagine that the State believes that ninety percent of the group does not condone administrative detention (in which case that State's utilization of detention will offend nine out of ten group members), and in the other it believes only sixty percent of the group does not condone administrative detention (i.e., that the procedure will only offend six of ten group members). Depending on its own beliefs regarding administrative detention, this difference may have an impact on the State's willingness to undertake the practice.

These differences can be measured in terms of their impacts on expected utility. Recall that the State may evaluate the procedure both positively as providing greater protection against terrorist activities, and also negatively as reducing civil liberties. Assume that only these two beliefs, along with a belief regarding the subjective norm, are relevant to the behavior. Assume further that beliefs about the behavior are held constant—our State prefers the use of administrative detention but does not hold this preference very strongly (as set forth above, +900 units). The certainty with which it holds beliefs about the norm will thus determine the State's attitude regarding the behavior. Consider the effect on utility in the situation when certainty of belief regarding the subjective norm drops from ninety-five to thirty, as set forth in Figures 4 and 5:

---

61. Rational choice scholars intuitively understand this attraction. See, e.g., Robert Cooter, *Do Good Laws Make Good Citizens? An Economic Analysis of Internalized Norms*, 86 VA. L. REV. 1577, 1592-93 (2000) ("Business, politics, love, and war cause people to form relationships with each other. These relationships create opportunities for mutual benefit from cooperation and also opportunities for people to exploit each other.").

62. JOHN C. TURNER ET AL., *REDISCOVERING THE SOCIAL GROUP: A SELF-CATEGORIZATION THEORY* 20 (Basil Blackwell Inc. 1987) ("[W]here people perceive, believe or expect to achieve mutual satisfaction from their association, they will tend to associate in a solitary fashion, to develop positive interpersonal attitudes and to influence each other's attitudes and behaviour on the basis of their power to satisfy needs for information and reward each other.").

Figure 4

Belief	Certainty	Evaluation	be
Will result in Sanction	+95	-10	-950
Protects against Terrorist Acts	+50	+50	+2500
Restricts Civil Liberties	+40	-40	-1600
		$A_o = \sum b_i e_i$	= -50

Figure 5

Belief	Certainty	Evaluation	be
Will result in Sanction	+30	-10	-300
Protects against Terrorist Acts	+50	+50	+2500
Restricts Civil Liberties	+40	-40	-1600
		$A_o = \sum b_i e_i$	= +600

Because the State's understanding of the uniformity with which a belief is held (+30 versus +95) impacts its estimation of normative sanction, the State will feel constrained to act by normative control in the latter case (because the utility is -50) and not constrained in the former (where the utility remains positive, +600).

Now, consider that our group member values group membership much more strongly. For example, contrast the situation of an emerging democracy with that of an established democratic regime. To the nascent State, approbation from and inclusion in the broader global community may be valued as more crucial to its international standing than would be the case for an accepted group member.<sup>63</sup> Thus, its negative evaluation of the normative consequences of acting out of step with group attitudes will be much more substantial (we assume an evaluation for this example of -50).

---

63. We recognize that other factors will also play a significant role in this process. For example, the anchoring and framing heuristics may significantly impact estimations of a nation's perception of the group norm. A nation that has, for many years, shown a strong proclivity for group membership, but that violates norms a few times, may thus do so with a certain amount of immunity. Continued violation, however, may clearly signal a change in the importance of group membership to the nation. See Alex Geisinger, *Are Norms Efficient? Pluralistic Ignorance, Heuristics, and the Use of Norms as Private Regulation*, 57 ALA. L. REV. 1, 19-24 (2005) (describing how heuristics may influence norm perception generally).

This evaluation will result in conformity regarding a much larger number of behaviors, as illustrated in Figure 6:

Figure 6

Belief	Certainty	Evaluation	be
Will result in Sanction	+30	-50	-1500
Protects against Terrorist Acts	+50	+50	+2500
Restricts Civil Liberties	+40	-40	-1600
		$A_o = \sum b_i e_i$	= -600

In such a case, even a small perceived certainty (+30) of sanction will result in conformity.

The more a State conforms to perceived group norms, the more likely other group members are to perceive that State to be strongly attracted to the group. The willingness of the above-mentioned emerging State to abide by international norms reflects its deep value of acceptance by the global community. This is particularly the case when States exhibit group conformity with less certain norms that are not universally held. This commitment to group membership acts as a strong signal of the emerging State's unwillingness to defect from cooperative endeavors with other group members.

The need-reinforcement addition to the basic rational choice model of behavior thus establishes a very particular view of groups and norm formation with the rational individual at its core. Norms arise only because rational individuals attain benefits from interacting with others and thus develop a free standing desire for others' acceptance. Individuals attempt to determine the preferences of the majority, and failure to act in accordance with the view of others negatively impacts one's perceived attractiveness to other group members. The higher one values group membership, the less likely one is to defect from cooperative endeavors.

The need-reinforcement principle provides a basis for including in the model of expressive effects of law a separate preference for esteem from others. The Reasoned-Action model describes this desire for esteem as the subjective norm. The desire for esteem will join other preferences that underlie a State's attitude to determine the utility of any particular regulated behavior. Such an understanding of norms allows us to model how international law and process may affect decisions regarding the utility of entering into and abiding by a particular legal regime.

## II. AN EXPRESSIVE MODEL OF RATIFICATION AND COMPLIANCE

According to the expressive model, international legal process can affect a State's understanding of the social norm through two distinct pathways. First, international process can change other States' preferences by providing information on objective reality. As the preferences of a number of States change, so too may norms and/or norm certainty. In cases where preferences change enough, a new norm may develop and become entrenched. Second, international process may directly communicate information on existing aggregate preference. That is, when normative belief is uncertain or not yet developed, international process can provide information on the beliefs of other States that will permit a State to develop its own understanding of the aggregate preferences of others. In such a case, norms haven't changed but certainty regarding the norm has.

### A. *International Processes*

Under the expressive model, international law and process are seen as means of attempting to change behavior by changing beliefs either about objective reality or about the subjective norm. Ratification of a treaty also impacts a number of beliefs about norms and thus also affects compliance. Let us consider each in order, using examples that isolate each of the different forces at work. We will start with the case of global warming.<sup>64</sup> The first element of the expressive model suggests that one way to change a State's desires regarding the regulation of the causes of global warming or compliance with a global warming treaty is simply to update its belief of objective reality. Consider the way in which over time science has changed estimations of the likelihood that global warming, with its attendant negative consequences, is occurring as the result of human-made changes to the natural environment. At some point, the certainty of harm will increase until it overcomes any reservations about ratifying or complying with a treaty. Take as an example a model of a State with only three salient beliefs regarding compliance with a global warming treaty. Figure 7 illustrates the salient beliefs held by that State during a period of early science in which there is little certainty that human-made materials change global climate.

---

64. As we will discuss *infra*, compliance with human rights treaties results primarily from the impact of international law on the subjective norm. However, international law may influence compliance through other expressive forces, such as by changing perceptions of objective reality that, in turn, change preferences for regulation. We turn to the example of environmental treaties in these sections to illustrate a variety of these expressive forces specifically because human rights treaties alone cannot be used to do so.



Figure 8 illustrates the salient beliefs held by that State during a period of later science when there is greater certainty that human-made materials change global climate.

**Figure 7**

Belief	Certainty	Evaluation	be
No regulation will cause sea levels to rise resulting in coastal region losses.	+20	-80	-1600
No regulation will increase skin cancers.	+15	-90	-1370
No regulation will benefit economy	+80	+90	+7200
		$A_o = \sum b_i e_i$	+4230

**Figure 8**

Belief	Certainty	Evaluation	be
No regulation will cause sea levels to rise resulting in coastal region losses.	+60	-80	-4800
No regulation will increase skin cancers.	+60	-90	-5400
No regulation will benefit economy	+80	+90	+7200
		$A_o = \sum b_i e_i$	-3000

Under this model, changing belief about objective reality results in a changed attitude toward regulation or compliance.

Of course, changing attitudinal belief can also change beliefs regarding the subjective norm. Norms, under our theory, simply reflect the aggregated preferences of members of the group of States to which that State belongs. As changes in understandings of objective reality result in increased preference for regulation of global warming, this in turn changes estimations of the likelihood of being sanctioned for failing to abide by the preferences of the majority of group members. At some point, the changed perception of group preference

will reach a tipping point and a new, pro-regulation norm will become entrenched.

International process may also provide information that is directly relevant to the subjective norm. Take as an example a newly formed Islamic democracy. Assume that this country's culture historically did not value women's social and political participation at a level equivalent to the beliefs currently held by the majority of democratic States. When considering whether to ratify the already existing Convention on the Elimination of All Forms of Discrimination against Women ("CEDAW"),<sup>65</sup> the Islamic democracy will infer from the CEDAW's existence a prevailing norm that favors women's equality.<sup>66</sup> To the extent that the Islamic democracy values membership in the global community of democratic States, it will feel increased pressure to ratify the CEDAW as a way to avoid social sanction. In this manner, the existence of treaties or other international regimes can inform directly the subjective norm.

Context plays a significant role in the way that ratification occurs. Given their dependence on scientific information, environmental treaties may be influenced to a large degree by changes in objective reality, like recognition that global warming is the result of man-made emissions. Trade regimes may be more dependent on reciprocal benefits or coercion, for example, conditioning a State's trade status on membership in the World Trade Organization or on its agreement not to counterfeit goods. Human rights regimes, on the other hand—which in the main are neither reciprocal nor influenced by changing scientific understandings—are relatively more dependent on direct normative influence; in other words, States ratify the CEDAW because of a general desire to be members of the international community rather than new information about the biological realities of women.

### *B. The Interaction of International and Domestic Processes*

An important issue in the developing compliance literature concerns how ratification of a treaty may influence domestic

---

65. G.A. Res. 34/180, at 193, U.N. GAOR, 34th Sess., Supp. No. 46, U.N. Doc. A/34/46 (Sept. 3, 1981).

66. An additional factor that may impact normative belief is the CEDAW's embodiment of a more universally held norm of equality. To the extent that a State believes equality to be a universally held norm, it will expect a large number of nations to be willing to enforce it. See Richard H. McAdams, *The Origin, Development, and Regulation of Social Norms*, 96 U. MICH. L. REV. 338 (1997) [hereinafter McAdams, *The Origin*].

lawmakers to regulate behavior.<sup>67</sup> That is, once a treaty has been ratified by a State, why and when does the State choose to conform to the treaty's terms? A simple answer is that States that ratify treaties have obligations to abide by them. This notion motivates claims that liberal States—that is, States that believe in the rule of law—are more likely to abide by their legal obligations than others.<sup>68</sup> However, this response does not address the more indirect way in which treaty ratification may permeate the “consciousness” of a State so as to affect everything from judicial decision making to domestic regulation and even the behavior of individual citizens. A developing body of international law scholarship asks the important question of how international norms may affect domestic understanding of an issue. This Section provides a more detailed analysis of how this mechanism works.<sup>69</sup>

Ratification of a treaty may influence domestic norms, which in turn can affect both citizen behavior and domestic policy. By demonstrating how ratification affects norms and preference, the expressive theory provides an understanding of the mechanisms by which this happens. Generally, ratification of a treaty may provide information on the aggregate preference of the majority in a liberal state where acts of the State's representatives (in this case the ratification of a treaty) are sometimes presumed to reflect the will of the majority.<sup>70</sup> To the extent this is the case, a treaty represents the

---

67. See, e.g., Goodman & Jinks, *Measuring the Effects*, *supra* note 6, at 173 (“[T]he incorporation of human rights norms [into domestic law] is a process; treaty law plays an important role in this process.”); Hathaway, *Do Human Rights Treaties*, *supra* note 5, at 1938-39 (discussing how “treaties play an important constraining role” in domestic law). The focus on domestic politics’ influence on compliance is often associated with the liberalist view. For an excellent discussion, see Hathaway, *Do Human Rights Treaties*, *supra* note 5, at 1952-55.

68. See, e.g., Kenneth W. Abbott, *International Relations Theory, International Law, and the Regime Governing Atrocities in Internal Conflicts*, 93 AM. J. INT’L L. 361, 366 (1999) (“[S]cholars are exploring whether liberal democratic states—with representative institutions and a commitment to the rule of law—are more amenable to legal relationships and arguments and more prone to comply with legal rules than states with different domestic regimes.”); Anne Marie Burley, *Law Among Liberal States: Liberal Internationalism and the Act of State Doctrine*, 92 COLUM. L. REV. 1907, 1920-21 (1992) (arguing that liberal States are more likely to act within the “zone of law” than others).

69. See, e.g., Jeffrey T. Checkel, *International Norms and Domestic Politics: Bridging the Rationalist-Constructivist Divide*, 3 EUR. J. INT’L REL. 473, 474 (1997) (noting that the domestic effect of international norms is to sometimes constrain and sometimes constitute); Andrew P. Cortell & James W. Davis, Jr., *Understanding the Domestic Impact of International Norms: A Research Agenda*, 2 INT’L STUD. REV. 65, 86 (2000) (“[T]he effects of an international norm cannot be understood independent of the norm’s salience in the domestic political discourse.”).

70. See also Robert E. Scott, *The Limits of Behavioral Theories of Law and Social Norms*, 86 VA. L. REV. 1603, 1647 (2000) (incorporating endogenous preference derived from social norms into rational choice theory). See generally Richard H. McAdams, *An Attitudinal Theory of Expressive Law*, 79 OR. L. REV. 339 (2000) [hereinafter McAdams, *An Attitudinal Theory*]

will of the electorate, and individuals will update their beliefs regarding majority preference accordingly. Ratification of a treaty may also inform individual preference which, in turn, further changes estimations of normative belief. As normative belief becomes more congruous with the treaty that was ratified, failure to abide by it will more likely result in social sanctions for individual behaviors contrary to the treaty's goals. In these ways, the certainty of social sanction for failure to abide by a norm of the international community becomes greater with the ratification of a treaty. To the extent that domestic policy in liberal States reflects the will of the public, changes in perceptions of majority preference will also influence domestic law making. Moreover, changed understanding of objective reality may also influence judicial decision making.

The following Subsections analyze how ratification affects normative belief and forces of domestic compliance. The first Subsection considers how ratification may directly impact the subjective norm of the ratifying State's citizens. Once we have demonstrated how ratification affects belief, we will discuss how and when changed belief translates into different market behavior and domestic regulation. The second Subsection considers indirect influences of ratification on the subjective norm.

### 1. Direct Influences on the Subjective Norm

There are a number of ways in which ratification can directly influence the subjective norms of citizens in a State that has ratified a treaty. Primary mechanisms again include providing information that changes individual preference, which may ultimately affect perceptions of majority belief, and providing information that leads one to infer an increased likelihood that the international norm will be socially enforced at a domestic level. We will examine each of these possibilities, starting with the latter.

Ratification of a treaty can provide information on what the majority of a State believes and thus serve to change or reinforce what citizens will sanction socially and, in the case of some democratic regimes, affect politicians' willingness to act. The mechanism by which ratification influences normative belief will be particularly influential in cases where there is great uncertainty regarding aggregate preference or objective reality. Let us consider the historical example of perceptions of ozone depletion in the United States up to and after

---

(presenting a theory that law changes behavior by signaling the underlying attitudes of a community).

ratification of the Montreal Protocol.<sup>71</sup> In 1986 great uncertainty still existed regarding the effects of human-made emissions on ozone.<sup>72</sup> America's ratification of the treaty, however, served as a coalescing moment that mobilized the population to act in support of the treaty's goals.<sup>73</sup> Indeed, just weeks after the United States ratified the Montreal Protocol, many communities began to ban products made with ozone depleting chemicals, "giving American industry a clear indication of the public mood."<sup>74</sup>

Expressive theory provides a means for describing this "ratification effect." Treaty ratification may change beliefs through a number of different inferential mechanisms. First, in States where treaties are subject to the approval of politically accountable branches before they can be ratified, ratification can carry with it information that a treaty's standards reflect the will of the majority. A person who is unsure of current majority preference but who believes that treaties will be ratified only when the majority approves of their standards will update his or her belief regarding the preferences of the majority as a result of ratification.<sup>75</sup>

The United States's ratification of the Montreal Protocol also provided a basis for an individual to increase his or her certainty of objective reality (in this case that ozone depletion is a significant problem) and affect preferences for regulation. The inferential process is similar to the one discussed earlier regarding how the willingness of other States to ratify may affect an individual State's belief. If one believes his or her State is generally unwilling to regulate behaviors subject to a treaty, but the State ratifies in any case, one may infer

---

71. The Montreal Protocol on Substances that Deplete the Ozone Layer, Sept. 16, 1987, 1522 U.N.T.S. 3, was ratified by the United States on April 21, 1988, and came into force on January 1, 1989.

72. See RICHARD BENEDICK, *OZONE DIPLOMACY: NEW DIRECTIONS IN SAFEGUARDING THE PLANET* 15 (Harvard Univ. Press 1998) (stating that "great uncertainties still remained" in 1986 in understanding the science of ozone).

73. See *id.* at 315 (describing how government efforts to inform their constituents about treaty obligations influenced politicians and private entities to act responsibly for the environment).

74. *Id.* at 102.

75. There is a concern regarding this mechanism. Many citizens simply do not believe that ratification correlates with majority beliefs. Public choice theory in particular suggests that special interest groups capture legislatures and thus that legislative acts do not reflect popular attitudes. Individuals that believe this to be the way legislation is created will not infer that legislation reflects the will of the majority. McAdams responds to this concern by suggesting that there are "weak" and "strong" forms of public interest theory and that, in some cases, inferences of majority influence on decision-making are still possible. See McAdams, *An Attitudinal Theory*, *supra* note 70, at 330-31. As our expressive theory clarifies, the strong or weak version of public choice theory is relevant to the effect of ratification on the subjective norm to the extent they are preexisting beliefs held by individuals. The weaker the belief in public choice theory, the more likely an individual is to believe that law correlates with majority opinion.

from ratification that the problem was worse than originally contemplated.<sup>76</sup> Indeed, ratification may well have provided the main impetus for a change in preferences, and ultimately norms, in the United States. As individuals inferred from United States ratification that the problem of ozone emissions was “real” or greater than they previously believed, they adjusted their preferences for regulation accordingly. As more individuals changed from a preference for no regulation to a preference for regulation, a tipping point was reached and a new norm entrenched.

We should note that our focus is on the effects of ratification on norms. This should not be taken to suggest that other elements of international lawmaking cannot affect normative belief. In the case of ozone, United States-led efforts mobilized much of the scientific community to create the evidence needed to link soundly CFCs to ozone depletion. The science alone, in cases where there is not competing science, could very well affect the preferences of those who learn of it through newspaper articles or other sources. The impact of international process on belief should not be underestimated. However, as the effect of the Montreal Protocol suggests, ratification itself is an important and distinct influence on domestic regulation and behavior.

We have established up to this point that treaty ratification can influence individual preferences, as well as beliefs about the preferences of others. Both of these forces affect the decisions of individuals within the marketplace. In the case of ozone, information linking CFCs to things such as Styrofoam food packaging and aerosol spray cans changed people’s preferences regarding the types of products they chose to use. As more people changed preferences regarding products, a norm in favor of products made without CFCs became entrenched, adding pressure to those who still preferred CFC-produced products to change their behavior or face social sanction. This substantially decreased demand for such products and cleared the way for new CFC-free technology.<sup>77</sup>

A change in norms would also likely affect the willingness of regulators to pass laws controlling the use of CFCs—at least when regulators see themselves as accountable to the populace. This, of course, highlights again the reasons why liberal regimes may be more greatly influenced by international process than others. Not only are

---

76. See Richard McAdams & Dhammika Dharmapala, *The Condorcet Jury Theorem and the Expressive Function of Law: A Theory of Informative Law*, 5 AM. L. & ECON. REV. 1 (2003).

77. See BENEDICK, *supra* note 72, at 104-05 (describing how ratification of the Montreal Protocol led to consumer pressure which in turn led to phase out of CFC products such as spray cans in England); *see also id.* at 310-11 (describing the “decisive” effects of consumer pressure on phase outs of products made with CFCs in England, Taiwan and Germany).

individuals in a liberal state more likely to believe that law reflects majority preferences and thus update their own normative belief, but politicians are also more likely to care about majority belief when they choose to regulate. Totalitarian and other non-democratic regimes feel less pressure from a change in citizen perceptions. This is not to suggest that what citizens perceive to be important has no impact in such regimes; the effects on citizens described above may also influence leaders. Similarly, to the extent that international process may affect preference and not just normative belief, the governing members of society may change their preferences regarding the need for regulation. Where the influence of citizen concern is omitted, however (such as in non-liberal regimes), the ability of ratification and process to effect change will decrease. The key to compliance in such cases is not the domestic impact on the citizenry, but the impact on government officials' preferences and their concern about the esteem from other members of international society.

Finally, we must consider how changing perceptions also influence judicial decision making through these mechanisms. A clear example is United States Supreme Court jurisprudence on the death penalty, determining the extent to which this punishment is "cruel and unusual" under the Eighth Amendment.<sup>78</sup> To make this assessment, the Justices reference "evolving standards of decency that mark the progress of a maturing society."<sup>79</sup> Abolishing the death penalty for convicted felons under the age of sixteen as offensive to "civilized standards of decency" in 1988, the Court noted that its conclusion was "consistent with the views that have been expressed by respected professional organizations, by other States that share our Anglo-American heritage, and by the leading members of the Western European community."<sup>80</sup> The following year the Court held that executing the mentally retarded could not categorically be said to violate either the Eighth Amendment or national norms.<sup>81</sup> However, because of the consistent American state movement away from this practice, as well as its almost complete absence internationally, in 2002 the Court banned those executions.<sup>82</sup> Last year the Court further proscribed the death penalty by prohibiting executions of juveniles under the age of eighteen, relying heavily on the "relevance of the views of the international community."<sup>83</sup> Among the evidence cited

---

78. For international parallels, see WILLIAM A. SCHABAS, *THE ABOLITION OF THE DEATH PENALTY IN INTERNATIONAL LAW* (Cambridge Univ. Press 2002).

79. *Trop v. Dulles*, 356 U.S. 86, 100-01 (1958).

80. *Thompson v. Oklahoma*, 487 U.S. 815, 830 (1988).

81. *Penry v. Lynaugh*, 492 U.S. 302, 334 (1989).

82. *Atkins v. Virginia*, 536 U.S. 304, 316 (2002).

83. *Roper v. Simmons*, 543 U.S. 551, 575 (2005) (internal citation omitted).

was the United Nations Convention on the Rights of the Child, which includes a ban on juvenile executions.<sup>84</sup> Because that treaty had been ratified by every State worldwide except Somalia and the United States, the Court found that “the United States now stands alone in a world that has turned its face against the juvenile death penalty.”<sup>85</sup>

The Supreme Court, in its death penalty jurisprudence, explicitly recognizes processes of expressive theory. In looking to the beliefs of other States, the Court responds to two main concerns. The first is the impact of jurisprudence on global esteem, which the Court explicitly recognizes when it notes that approving the death penalty would leave the United States “standing alone” in the community of States. The Court also clearly demonstrates its understanding that international norms and treaties may influence domestic jurisprudence by influencing our domestic beliefs regarding the underlying issues. The exogenous influence is a product of the United States’ self-identification as a culturally Western State. Consider the simple syllogism behind the Supreme Court’s analysis of the death penalty: (1) the United States is a civilized State with strong ties to Western States and their traditions, (2) both civilized States and culturally Western States uniformly believe the death penalty should not be applied in certain situations, (3) the consistently held beliefs of these States should be instructive to the United States’ own view of the death penalty in these situations. This line of reasoning recognizes that one source of information on objective reality is the beliefs of other, similarly-situated States. Put simply, the Court reasons: “They are like us and they don’t apply the death penalty, we might infer from this that there is something uncivil about its application.” Domestic belief is not changed because of specific information regarding the particular harms of the death penalty (for instance, if empirical evidence demonstrated that the existence of the death penalty increased the likelihood of homicide because criminals believed they had no incentive to refrain from egregious violence). Rather, belief may be changed because of inferences drawn from others’ changed beliefs. This dynamic interplay between group members provides the basis for international norms to effect domestic norms.

We have to this point identified the processes by which international regimes may expressly influence domestic law and behavior. International process and ratification can influence domestic regulation as well as market behavior through a number of different

---

84. *Id.* at 576. United Nations Convention on the Rights of the Child, GA. Res. 44/25, at 161, U.N. GAOR, 44th Sess., Supp. No. 49, U.N. Doc. A/44/49 (Nov. 20, 1989).

85. *Roper*, 543 U.S. at 577 (internal citations omitted).



mechanisms. Where process produces information, that information may influence preference and in turn perceptions of the subjective norm. Ratification in certain conditions may have similar effects. For instance, ratification can also carry information on what behaviors others prefer or are willing to socially sanction. There are thus a wide variety of ways in which treaty-making and ratification can influence domestic processes.

These processes do not work in a vacuum; context must also be considered. A number of factors may also increase or decrease the domestic effects of ratification. If a State develops domestic law to comply with a treaty's standards, that law may, in turn, have further expressive impact. Similarly, the pre-existing beliefs of the populous may significantly influence the domestic impact of international process. We have already discussed how preexisting belief may influence perceptions of objective reality. Preexisting belief may also affect beliefs about the subjective norm. If, for example, a State's citizens believed the State's leaders were ratifying a treaty only due to international pressure and in spite of public preference, ratification would likely have no effect on understanding of majority preference. In this way, each situation of international regulation presents its own unique set of factors to which the tools of expressive theory can be applied in order to determine their likely domestic affect.

## 2. Indirect Influences on Normative Belief

Indirect influences can also affect a State's normative belief and thereby change its actions. Individuals and entities without formal positions in the political sphere, such as persons with recognized moral status (e.g., an imprisoned Nelson Mandela),<sup>86</sup> celebrities from the entertainment world (for instance, French actress Brigitte Bardot, a strong advocate of animal rights),<sup>87</sup> and media outlets,<sup>88</sup> can impel States to honor their international obligations by influencing estimations of the likelihood of sanctions. Among indirect actors, non-governmental organizations ("NGOs") often play a key role in promoting international law observance by influencing the

---

86. See NELSON MANDELA, *LONG WALK TO FREEDOM* (Little, Brown & Co. 1994). Put another way, certain individuals who are representative of the moral point of view can influence the perceptions of what behaviors are preferred by "moral" nations. See *infra* note 118 (describing the representativeness heuristic and its impact on norm perception).

87. BARNETT SINGER, *BRIGITTE BARDOT: A BIOGRAPHY* (McFarland & Co. 2006). High visibility increases the availability of information on a particular point of view, thus serving to influence perceptions of others' preferences. See *infra* note 119 (describing the availability heuristic and its impact on norm perception).

88. For a detailed description within one national context, see Benjamin L. Liebman, *Watchdog or Demagogue? The Media in the Chinese Legal System*, 105 COLUM. L. REV. 1 (2005).

subjective norm. This is especially true in the context of human rights treaties.<sup>89</sup>

A central function of NGOs is investigating international human rights violations and disseminating information on these abuses. NGO-issued human rights reports are usually framed against a State's previous ratification of one or more human rights conventions, and point to particular acts as violating those agreements.<sup>90</sup> Hence, NGOs use the historical fact of a State's prior agreement to abide by international norms, and its present desire for esteem from other States also bound by those accords, to shame a non-complying State into altering its conduct.<sup>91</sup> One commentator notes that this tactic is "particularly effective where a government is sensitive to international public opinion,"<sup>92</sup> but this understates the point. It is precisely because States value esteem from the global community that they are compelled, through the need-reinforcement principle, to alter their actions and thereby avoid the disapprobation of their peers. NGOs are especially well placed to alter a State's normative belief because their main currency is disseminating information that can erode a State's global reputation. Bringing pressure to bear through negative publicity leverages non-complying States to alter their behavior and honor their international obligations.<sup>93</sup>

To illustrate how an NGO can change a State's estimation of the subjective norm, consider the actions of Mental Disability Rights International ("MDRI"), an NGO promoting "international oversight"

---

89. See generally NON-STATE ACTORS AND HUMAN RIGHTS (Philip Alston ed., 2005) (discussing different approaches that might create international accountability for non-State actors). Although we focus on NGOs in this Article, we note that the contributors to this collection correctly point out and explain the role of other non-State actors, for example multinational corporations and international agencies, in enforcing human rights regimes.

90. The mandate of the global NGO Human Rights Watch, for example, includes documenting and reporting abuses that violate international instruments prohibiting torture and the subordination of women. An overview of that organization's activities is available at Human Rights Watch, <http://www.hrw.org> (last visited Feb. 15, 2006).

91. "The pointed finger of shame . . . has caused executions to be stayed, death sentences to be commuted, torture to be stopped, prison conditions to be ameliorated, prisoners to be released, and more attention to be paid to the fundamental rights of many citizens." David S. Weissbrodt, *The Role of International Nongovernmental Organizations in the Implementation of Human Rights*, 12 TEX. INT'L L.J. 293, 298 (1977).

92. Laurie S. Wiseberg, *Protecting Human Rights Activists and NGOs: What More Can Be Done?*, 13 HUM. RTS. Q. 525, 531 (1991).

93. See, e.g., Michael H. Posner & Candy Whittome, *The Status of Human Rights NGOs*, 25 COLUM. HUM. RTS. L. REV. 269, 272 (1994) (noting that NGOs "have helped to transform the debate from a restrained diplomatic discourse among governments to a more urgent and real confrontation of ongoing and serious human rights crises").

on the human rights of persons with intellectual disabilities.<sup>94</sup> After a two-year investigation, MDRI released a detailed report and press release of human rights abuses against mentally disabled children and adults in Hungary's national mental health system.<sup>95</sup> Among the findings of MDRI's investigators were that Hungarian children with disabilities were excluded from schools and placed for the rest of their lives in institutions; that adult residents of "social care" (i.e., nursing) homes, psychiatric institutions and general hospital psychiatric wards were subjected to lifelong inhuman and degrading treatment; and that among the "care" given to mentally disabled persons was being "placed into cages for unlimited time periods without supervision."<sup>96</sup> MDRI asserted that these actions violated established human rights covered by the International Covenant on Civil and Political Rights,<sup>97</sup> the European Convention on Human Rights,<sup>98</sup> and the International Covenant on Economic, Social and Cultural Rights.<sup>99</sup> In response to the international negative publicity generated by MDRI's report,<sup>100</sup> Hungary promulgated new disability rights legislation, and established a human rights ombudsman system to protect institutionalized persons.<sup>101</sup>

Hungary's reconsideration of the human rights of persons with mental disabilities in state-run institutions following MDRI's allegations of human rights abuses fit neatly within the need-reinforcement framework. At the time of MDRI's report, some eight years after its political emancipation from the former Soviet Union,

---

94. An overview is provided in the organization's website, which is available at Mental Disability Rights International, <http://www.mdri.org> (last visited Feb. 15, 2006). What follows is drawn from press releases, human rights reports, and sources posted therein.

95. MDRI's investigators visited and interviewed people in both institutional and "social care" (nursing home) settings. See ERIC ROSENTHAL ET AL., HUMAN RIGHTS AND MENTAL HEALTH: HUNGARY (1997), <http://www.mdri.org/pdf/Hungary.pdf>.

96. *Id.* at 16-18.

97. G.A. Res. 2200A, at 52, U.N. GAOR, 21st Sess., Supp. No. 16, U.N. Doc. A/6316 (Dec. 19, 1966).

98. European Convention for the Protection of Human Rights and Fundamental Freedoms, 213 U.N.T.S. 222 (Sept. 3, 1953).

99. G.A. Res. 2200A, at 49, U.N. GAOR, 21st Sess., Supp. No. 16, U.N. Doc. A/6316 (Dec. 16, 1966). Curiously, MDRI might also have asserted violations of the United Nations Convention on the Rights of the Child, G.A. Res. 44/25, at 166, U.N. GAOR, 44th Sess., Supp. No. 49, U.N. Doc. A/44/49 (Nov. 20, 1989), as well as the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, G.A. Res. 39/46, at 197, U.N. GAOR, 39th Sess., Supp. No. 51, U.N. Doc. A/39/51 (Dec. 10, 1984).

100. In contrast, widespread publicity within Hungary had almost no effect. See Gabor Gombos et al., *Hungary: The Social Care Home Report*, 21 N.Y.L. SCH. J. INT'L & COMP. L. 361, 363-69 (2002) (describing the pace of progress in Hungary's mental health care institutions following release of MDRI's report).

101. See HUMAN RIGHTS AND MENTAL HEALTH: HUNGARY, *supra* note 95, at 6-7 (describing new disability rights legislation promulgated by Hungary).

Hungary aspired to join the Western community of States and sought their approbation. The report and the publicity it engendered produced reputational damage and esteem loss for Hungary from the United States and the European Union, which in turn caused Hungary to recant its policies and alter its behavior. In other words, as an NGO, MDRI indirectly affected Hungary's estimations of utility by changing that State's perception of the likelihood that failure to abide by an international norm would result in a loss of esteem in the eyes of other countries. In order to counter the effects of the release of this information, Hungary took action that remedied the esteem loss.<sup>102</sup>

### III. CLARIFICATIONS AND IMPLICATIONS OF THE EXPRESSIVE MODEL

As we have just discussed, the expressive theory provides a detailed model of how international process affects the willingness of States to ratify and comply with treaties. One other strength of the theory is its ability to explain, at least in part, a number of issues where international law scholars either find themselves at odds with one another or where there is simply no reliable understanding of a particular State's action. These include providing a way of harmonizing divergent viewpoints on the ability of international regimes to construct the identities of States, whether the social force is different than traditionally understood forces of coercion and persuasion, how norms become internalized and what information signals may carry. This Part will apply expressive theory toward a better understanding of these and other issues.

#### *A. Harmonizing Rationalism and Constructivism*

Discussion of how international law affects behavior generally divides into two schools of thought: the rationalists, who emphasize military-economic power, and the constructivists, who focus on norms and global ideational structure.<sup>103</sup> The rationalist/constructivist divide also resonates within arguments regarding the primary means for achieving compliance among States. Accordingly, rationalists believe

---

102. MDRI has followed similar strategies in other emerging countries seeking global esteem. See generally Eric Rosenthal, *The Application of International Human Rights Law to Institutional Mental Disability Law*, 21 N.Y.L. SCH. J. INT'L & COMP. L. 387 (2002) (transcript of remarks by Eric Rosenthal, MDRI's founder and executive director). For additional, country-specific reports, see MDRI's website, *supra* note 94.

103. Goodman & Jinks, *How to Influence States*, *supra* note 6, at 632-33; see also Koh, *Internationalization*, *supra* note 2, at 976 (discussing how the dialectic between those who take interests as a given and those who take them as constructed developed).

that coercion is the force by which compliance is created, while constructivists emphasize persuasion.<sup>104</sup>

The rationalist/constructionalist divide has served to straight-jacket efforts to consider in depth the force of social processes on compliance. Lack of an independent model of how social forces work hampers discussion and design of international regimes. As Goodman and Jinks have observed, it is not that "international legal scholarship has completely failed to identify aspects" of how social forces affect state behavior. Rather, the mechanism by which this happens is "poorly understood . . . and often conflated (or confused) with other constructivist mechanisms such as persuasion."<sup>105</sup> Put simply, without a better understanding of how normative compliance forces operate, we cannot form an accurate account of how best to achieve compliance in any given situation.<sup>106</sup> The expressive theory provides substantial insights into this lacuna. By identifying the framework and microprocesses of normative influence, the expressive theory provides a means for discussing the differences between normative and other influences on State behavior. Further, the expressive theory suggests when the different mechanisms of change are most effective.

Our expressive theory is primarily a rationalist understanding of normative behavior. It is based on the notion that rational States have preferences to be "esteemed members" of international society due to the increased benefits that come from cooperation. However, the expressive theory diverges from purely rational actor methodology in two significant ways. First, it does not concur with the rational actor determination that coercion is the force by which compliance is created. Second, because expressive theory draws on the desire for esteem rather than on coercion as the motivating factor, it also recognizes that international processes have constructive effects.

Constructivists argue that interests are socially constructed by commonly held principles, identities, behavior norms, or shared terms

---

104. Goodman & Jinks, *How to Influence States*, *supra* note 6, at 633-34; Koh, *Internationalization*, *supra* note 2, at 976. For a discussion of these two schools, their assumptions and limitations, see Hathaway, *Do Human Rights Treaties*, *supra* note 5, at 1944-62.

105. Goodman & Jinks, *How to Influence States*, *supra* note 6, at 626-27; *see also* Koh, *Internationalization*, *supra* note 2, at 980-81; Alvarez, *supra* note 7, at 970 (interpreting Goodman & Jinks's concept of acculturation).

106. *See also* Dinah Shelton, *Compliance with International Human Rights Soft Law*, in *INTERNATIONAL COMPLIANCE WITH NONBINDING ACCORDS* 119, 131-32 (Edith Brown Weiss ed., 1997) (decrying adequate methodologies for adequately examining compliance); Beth A. Simmons, *International Law and International Relations: Scholarship at the Intersection of Principles and Politics*, 95 *AM. SOC'Y INT'L L. PROC.* 271, 273-78 (2001) (lamenting the dearth and depth of interdisciplinary examinations of central issues, including compliance).

of discourse.<sup>107</sup> This constructivist school also shares ties to the English “international society” school because both believe that States obey international rules for reasons beyond mere calculations of how compliance or noncompliance will affect their interests.<sup>108</sup> They believe instead that State decisions are influenced by the pull of international society, itself a force that derives from States’ long-term interests in the maintenance of international community.<sup>109</sup>

The need-reinforcement component of expressive theory provides the link between the constructivist and rationalist approaches. By harmonizing theories from these distinct schools of thought, it provides a more comprehensive framework for understanding and predicting State behavior. Need-reinforcement demonstrates how rational self-interest can lead to a separate desire (or preference) to maintain relations with other States. Like the rationalists, need-reinforcement describes the desire to be members of international society as a result of the benefits these States will receive through cooperation. This desire, however, exists separately from other preferences—it is a preference for esteem and not a preference for material or economic benefit—and can serve as a powerful influence on the willingness of States to conform to law.

Similarly, the need-reinforcement principle provides a better understanding of groups and group discourse. As we have seen in our earlier discussion of need reinforcement, the key to social influence is that an individual identify with a particular group. Further, that group membership derives from cooperative benefit borne of frequent and long-term interaction. At some point, group membership attracts one to other members of the group generally. Over time, the group will develop commonly-held norms. Behaviors for which clearly developed norms do not exist, however, would be open to group “negotiation.”

We start our discussion of how interests are constructed with the premise that not all norms are clearly developed or entrenched. In many cases, States simply may not have perfect (or even accurate)

---

107. See Koh, *Why Do Nations*, *supra* note 3, at 2633.

108. *Id.* at 2634 (“The predominantly American constructivist school has close familial ties to the English ‘international society’ school of Grotian heritage.”). Koh cites Andrew Hurrell, John Vincent, Barry Buzan, Gerritt Gong, Richard Little, and Michael Donelan as modern scholars who work in this vein. *Id.* at 2634 n.181.

109. *E.g.*, CHAYES & CHAYES, *supra* note 1, at 27-28; Martti Koskenniemi, *The Pull of the Mainstream*, 88 MICH. L. REV. 1947 (1990); John K. Setear, *An Iterative Perspective on Treaties: A Synthesis of International Relations Theory and International Law*, 37 HARV. INT’L L.J. 139 (1996); see also Tom R. Tyler, *Compliance with Intellectual Property Laws: A Psychological Perspective*, 29 N.Y.U. J. INT’L L. & POL. 219, 220-21 (1997) (advancing a psychologically based theory of compliance).

information on the preferences of others.<sup>110</sup> This normative misperception is often referred to as pluralistic ignorance, a condition in which nearly all members of a group believe that virtually all other group members support certain norms, even as these same individuals privately reject that behavior.<sup>111</sup> A renaissance of recent work in the field has demonstrated the phenomenon's existence in a wide variety of group contexts,<sup>112</sup> although the phenomenon has been understood for well over seventy years.<sup>113</sup> Although the reasons for the existence of pluralistic ignorance have not yet been fully developed, this wide-ranging experimental work demonstrates that a number of factors, including lack of information and misinformation regarding other States' preferences, may lead to uncertainty regarding norms in a wide-variety of contexts.<sup>114</sup>

When norms are uncertain, the process of international law making can serve to construct normative beliefs. Let us consider an example that will help identify some of these constructive mechanisms. Assume five States, in geographic proximity to each other and with strong diplomatic and economic ties, are experiencing an influx of immigrants from nearby States. Assume further that the

---

110. See McAdams, *An Attitudinal Theory*, *supra* note 70, at 355-56 (describing the variety of ways in which individuals may lack perfect information on the approval patterns of others).

111. *Id.* at 356-57.

112. See generally David Hines et al., *Pluralistic Ignorance and Health Risk Behaviors: Do College Students Misperceive Social Approval for Risky Behaviors on Campus and in Media?*, 32 J. APPLIED SOC. PSYCHOL. 2621, 2622 (2002) (stating that pluralistic ignorance is a pervasive feature of social life: "It has been found to characterize the dynamics of social situations . . . social groups . . . and social movements . . ."). For specific discussions of pluralistic ignorance in a variety of groups see Cristina Bicchieri & Yoshitaka Fukui, *The Great Illusion: Ignorance, Informational Cascades, and the Persistence of Unpopular Norms*, in EXPERIENCE, REALITY, AND SCIENTIFIC EXPLANATION 89, 93 (Maria Carla Galavotti & A. Pagnini eds., 1999) (citations omitted) (gang members and prison guards).

113. For example, as far back as the 1930s Richard Schanck demonstrated how members of the Methodist Church in Elm Hollow were almost unanimous in outwardly supporting norms against card playing, drinking alcohol, and smoking while privately undertaking these behaviors.

114. While a comprehensive understanding of pluralistic ignorance does not yet exist, those who study the phenomenon generally agree that it is based on a propensity of individuals to perceive their motivation for acting as different from others. See Deborah A. Prentice & Dale T. Miller, *Pluralistic Ignorance and Alcohol Use on Campus: Some Consequences of Misperceiving the Social Norm*, 64 J. PERSONALITY & SOC. PSYCHOL. 243, 244 (1993). In particular, social psychologists have demonstrated that an individual may understand his or her own actions to be normatively controlled but perceive others to act out of preference or true belief. See Hines et. al., *supra* note 112; see also Deborah A. Prentice & Dale T. Miller, *Pluralistic Ignorance and the Perpetuation of Social Norms by Unwitting Actors*, in 28 ADVANCES IN EXPERIMENTAL SOC. PSYCHOL. 161, 162 (Mark P. Zanna ed., 1996). For a discussion of the self/other difference and its sources, see Bicchieri & Fukui, *supra* note 112, at 97; see also Dale T. Miller & Deborah A. Prentice, *Collective Errors and Errors About the Collective*, 20 PERSONALITY & SOC. PSYCHOL. BULL. 541, 544-45 (1994) (suggesting the bias is based on the cultural propensity to underestimate the power of social motives to influence behavior).

five States are considering whether to enter into an agreement that provides immigrants with certain basic rights (say, health care and other financially costly social benefits) and that, at this point, three of the five States generally disfavor providing rights to immigrants, and two are slightly in favor of doing so.

During the process of negotiating an agreement, the parties will reveal their preferences more clearly. After a general discussion of the topic, an observer with good information would likely conclude that there exists a norm against providing rights to immigrants in this region, albeit a weak one. The negotiation, however, provides an opportunity for States to influence both the preferences and normative belief of others. A pro-treaty State may attempt to change other group members' perceptions of objective reality through the provision of information, for instance, by explaining how providing these rights will incentivize workers with technical skills to emigrate and fill a need in the workforce. The pro-treaty State may also attempt to change the reticent States' perceptions of the subjective norm. It may appeal to more broadly held norms such as "equality" or "fairness" as a means of either decreasing estimations of the likelihood of being sanctioned by the other States, or perhaps of attempting to convince other States that a pro-immigration norm exists within the group.<sup>115</sup>

In a similar vein, the pro-treaty State may appeal to well established and more specific norms in the global community to which the five States belong. If the State can point to a generally held universal norm of fair treatment of immigrants, this too can influence expectations of social sanction. This latter point recognizes that normative influence can transcend the boundaries of the specific group of States that are negotiating the treaty. States identify themselves as members of more than one group. They may find the influence of certain States with whom they share strong bonds to be greater than others to whom they are more loosely connected because they evaluate the esteem from these different group members differently. Such loose connections, however, still exert some degree of normative force based on the desire for esteem from these other group members. In this sense, group identity plays a key role in the influence of normative belief, and the beliefs of all States with whom a State identifies become relevant to its analysis of the subjective norm.<sup>116</sup>

---

115. McAdams has described such a phenomenon, explaining how certain specific behaviors may be made normative by connecting them to more universally held and generalized norms. McAdams, *The Origin*, *supra* note 66, at 383.

116. In this way a large block of nations, for example, Western Democracies, may exert influence on matters between other group members to which they are not even part. See generally Diane M. Mackie & Sarah Queller, *The Impact of Group Membership on Persuasion: Revisiting "Who Says What to Whom with What Effect?"*, in ATTITUDES, BEHAVIOR, AND SOCIAL



Finally, a number of heuristics<sup>117</sup> may further influence normative perception. States (as well as NGOs) may attempt to use States identified as representatives of the group to affect the norm estimations of other group members.<sup>118</sup> They may also attempt to make the information that supports their normative claims more available in order to skew norm perception.<sup>119</sup> If these forces change the intent of even one of the three reticent group members, then a tipping point is reached and a weak pro-treaty norm will be established. Expressive mechanisms that rely on self-identification, along with other efforts to make aggregate group preference more certain, can be used to influence the interests of rational States. Indeed, in the case of treaties that are dependent on the normative mechanism, this force will be the primary, if not the only, means by which international processes may proceed. This example makes clear that interests of others within the group or a different group with which the States identify can be used to affect the way interests are “constructed” through the commonly held principles, identities, behavioral norms and terms of discourse suggested by Koh.

In sum, the expressive theory recognizes that interest theory and constructive theory can be squared. While States maintain preferences for the esteem of others, given uncertain and constantly changing normative belief, established interests rarely exist at the outset of international discussions. Rather, discussions between States play the role of influencing normative belief and, in turn, the interests of States.

### *B. Distinguishing Persuasion, Reputation, and Coercion*

Failure to understand more clearly the differences between forces of coercion, persuasion and reputation hamper efforts at regime

---

CONTEXT: THE ROLE OF NORMS AND GROUP MEMBERSHIP, *supra* note 29, at 135 (providing empirical evidence of and discussing the importance of approval from other group members).

117. Heuristics are cognitive shortcuts that allow individuals to meaningfully process probabilistic information. See Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051, 1127-34 (2000) (describing the effect of social norms on deviations from self-interest).

118. The representativeness heuristic “refers to the tendency of actors to ignore base rates and overestimate the correlation between what something appears to be and what something actually is.” *Id.* at 1086; see also Dale T. Miller et. al., *Pluralistic Ignorance and Inconsistency Between Private Attitudes and Public Behavior*, in ATTITUDES, BEHAVIOR AND SOCIAL CONTEXT: THE ROLE OF NORMS AND GROUP MEMBERSHIP, *supra* note 29, at 95, 103-12 (discussing the exaggerated normative influence of those who embody the norms of a particular group).

119. Availability refers to the tendency to overestimate the likelihood of certain outcomes due to the fact that information regarding that outcome is more “vivid, well-publicized, or more prevalent among a particular actor’s friends and acquaintances.” Korobkin & Ulen, *supra* note 117, at 1087-88.

design and make identification of the most effective compliance tools unattainable. The expressive model provides a basic understanding of when a state is persuading, using reputational measures, or coercing compliance. While there is some overlap, the expressive model treats persuasion, reputation, and coercion as primarily distinct forces.

Consider first the distinction between persuasion and reputation. According to the expressive model, persuasion occurs when an actor attempts to use information regarding objective reality to revise estimates of certainty or evaluation of non-normative preferences, while normative or reputational influence occurs when information on the subjective norm is communicated. Consider the beliefs of the State in Figures 7 and 8 regarding global warming. Over time, the science regarding global warming changes the State's estimation of the likelihood that seas will rise, destroying a portion of its coastline and inland areas, and increasing the number of cancers. At some point, the certainty of these harms increases, in turn changing a preference from non-compliance to compliance. This is true persuasion: the information changes the State's beliefs to the point that it now prefers to comply. As discussed above, persuasion also affects normative belief; thus, no act is purely persuasive. However, to the extent that information acts directly to influence beliefs about objective reality, it would be fair to characterize the normative influence of this behavior as an indirect—and thus secondary—benefit. On the other hand, when efforts are focused directly on changing beliefs regarding the subjective norm, those efforts are primarily reputational.

Now consider the difference between reputation and coercion. The key to this distinction is that reputation works on the subjective norm while coercion implicates preference for factors beyond the scope of the treaty. For example, specific aid money or restraint from the use of military force may be considered coercive because each operates on preferences that lie outside of a treaty regime. By contrast, reputational forces work by their appeal to the general benefit one derives from being a group member.

Note that in this respect, coercion is not like persuasion in that the latter relates to the goals of the international regime and provides a means for explaining how the regime itself will further preference. Coercion recognizes that persuasive and reputational forces are not sufficient to create compliance. Rather, there are other benefits or detriments necessary to achieve compliance goals. In this sense, reputational forces exhibit the characteristics of both coercion and persuasion. They are coercive to the extent that they are outside the attitude of a State; they are persuasive because they reflect an existing preference for esteem. Reputational forces are distinct from

coercion and persuasion, however, in that they act on a particular preference—the preference for the esteem of other States.

Another dividing point between coercion and persuasion is autonomy. In the event of coercion, a State's autonomy is violated because it is brought to comply with a position that it would otherwise oppose. Persuasion may likewise bring a State to heed a regime dynamic that under other circumstances it would resist, but the State's autonomy is maintained because it "voluntarily" agreed to this otherwise undesired regime. In real world terms, these categories may devolve down to the proverbial stick and carrot, even if the end result looks much the same.<sup>120</sup>

### C. Explaining Internalization

The expressive model also helps us to understand exactly how the rules of a particular treaty become internalized. Many scholars recognize that voluntary obedience arising from internalization of a norm, rather than coerced compliance, is the preferred enforcement mechanism.<sup>121</sup> The question that remains, however, is how norms become internalized. To date, current understanding of internalization within international law scholarship has been limited.<sup>122</sup> Generally, internalization is equated with a dynamic process of identification, repetition and feedback. Harold Koh, for example, describes the process as starting with the identification of a norm as a result of an interaction between transnational actors.<sup>123</sup> Once the norm is announced, certain transnational actors seek to bind other States to

---

120. In defining this taxonomy, we expressly reject purely existential notions of volition. For example, we disallow claims that poor women without employment prospects and in dire economic straights "voluntarily" engaged in prostitution or other illegal activity.

121. CHAYES & CHAYES, *supra* note 1, are emblematic of this assertion. For additional, and varied examples, see Robert C. Bird, *Procedural Challenges to Environmental Regulation of Space Debris*, 40 AM. BUS. L.J. 635, 645-49 (2003) (discussing the promotion of cooperation rather than sanctions); George W. Downs et al., *The Transformational Model of International Regime Design: Triumph of Hope or Experience?*, 38 COLUM. J. TRANSNAT'L L. 465, 468 (2000) (discussing how empirical evidence of horizontal cooperation "carrots" shows it as more efficacious than vertical coercive "sticks"); Tom Ginsburg & Richard H. McAdams, *Adjudicating in Anarchy: An Expressive Theory of International Dispute Resolution*, 45 WM. & MARY L. REV. 1229, 1243-88 (2004) (providing a game theory model to demonstrate the proclivity of States to cooperate due to inter-connected relationships); Guzman, *supra* note 3, at 1849 (finding that States are motivated by cooperative gains).

122. See Posner, *supra* note 13, at 824 (arguing that, while constructivist scholars rely heavily on the notion of internalization, they have yet to develop a theory of how internalization happens).

123. See Koh, *Why Do Nations*, *supra* note 3, at 2646 (internal citations omitted) (describing three phases of processes of transnational actors).

“obey the [norm] as part of [their] internal value set.”<sup>124</sup> According to Koh:

Such a transnational legal process is normative, dynamic, and constitutive. The transaction generates a legal rule, which will guide future transnational interactions between the parties; future transactions will further internalize those norms; and eventually, repeated participation in the process will help to reconstitute the interests and even the identities of the participants in the process.<sup>125</sup>

This model, however, confuses compliance that results from a desire for esteem or from habit with the process of internalization. A simple thought experiment highlights the difference. Imagine a group of States with a well-entrenched norm that education is a primary basic human right.<sup>126</sup> Imagine further that one member of the group is an agricultural country with poor economic conditions, where the basic goal of State spending is subsistence and public health. Education is not accorded as high a preference in such a country, but it nevertheless feels pressure from the well-established norm to provide education to its populace.

Now assume that the other group members have lost their ability to determine whether the agricultural State is providing education so that there are no more reputational effects tied to compliance. At some point, one would expect the practice of providing education to evolve and the preference of using resources to other ends to assert itself.<sup>127</sup> Long-term interaction with other States regarding the well-established norm has not resulted in that norm somehow permeating the consciousness of the agricultural State; internalization has not occurred. Rather, internalization occurs in circumstances where, if the normative factors are removed, a State still believes it is in its best interest to follow a particular behavior.<sup>128</sup> The State, in Koh's terms, obeys because of its own internal values. In that case, the State is no longer acting out of regard for external social sanctions (opprobrium by fellow States for not providing public education), but it is acting out of preference (because it has come to regard positively the provision of public education). Nor is it likely that simple repetition of a behavior (providing public education) will affect its internalization.

---

124. *Id.*

125. *Id.*

126. As indeed, they have through the Convention on the Rights of the Child, G.A. Res. 44/25 at 166, U.N. GAOR, 44th Sess., Supp. No. 49, U.N. Doc. A/44/49 (Nov. 20, 1989).

127. See Korobkin & Ulen, *supra* note 117, at 1113-15 (noting that habits can be rational in the global sense but recognizing that past behavior does not affect the preference structure of an actor).

128. See Robert E. Scott, *The Limits of Behavioral Theories of Law and Social Norms*, 86 VA. L. REV. 1603, 1626 n.51 (2000) (noting that internalization must be the result of changed preference); Geisinger, *supra* note 17, at 66 (arguing that internalization can only be equated with a preference change).

These beliefs are more likely to change due to changes in information about objective reality.<sup>129</sup>

This is not to imply that the process outlined by Koh does not influence behavior (or that his contributions to the field have been anything less than seminal). What we point out is that when these processes act, they do so through the mechanism of increasing norm certainty suggested by the expressive model, rather than internalization as suggested by Koh. Indeed, the above quotation provides wonderful insight into how complex interaction can construct and develop normative belief. But consider Koh's example in a new light; that is, let us re-write his appraisal in the context of the forces that impact normative belief. At the beginning of the process, the norm relevant to a situation is unknown or, at most, uncertain. Announcement of the norm begins the move toward increased certainty. As other States are guided by the norm, certainty that a particular behavior is norm-congruent increases, with a corresponding increase in the esteem a State would expect from acting in accordance with the norm. This leads more States to act in accordance with the norm. Over time, this norm cascade continues until a point where the norm becomes entrenched in the fabric of international society.

This reinvention of Koh's intrepid insights helps clarify how State desire for esteem can be harnessed by international legal process. Legal process provides not just focal points for cooperation, but also an iterative process of norm development and entrenchment that carries with it strong influence on the behavior of States.

#### *D. Explaining Signaling*

The expressive theory also adds insights to discussions of how States may signal each other through treaty ratification and compliance. Signaling theory takes as its starting point the presumption that States communicate their cooperative nature to others by undertaking certain behaviors that reflect their willingness to forego immediate benefit for long-term gain.<sup>130</sup> States that undertake the most expensive signals demonstrate their low discount rates; that is, they demonstrate the higher value they place on future

---

129. There are a number of real-life examples of this phenomenon. Eastern European nations that were previously members of the Soviet Union, for example, now speak their native languages instead of Russian, and have also returned to their own native faiths and national identities. Much the same can be said for the situation in Northern Ireland.

130. For a more complete discussion, see POSNER, *supra* note 46, at 18-22 (providing the foundational exegesis of signaling theory in law); Moore, *supra* note 8, at 883-85 (adapting Posner's theory to the context of human rights).

payoffs than States that don't undertake such behaviors.<sup>131</sup> These low discount rates translate into those States being better cooperative partners because it is less likely that a state that values long-term gains will defect from cooperative relationships for short-term benefits.<sup>132</sup>

A signaling theory of international law proceeds to consider the types of signals that States create through ratification and compliance behavior. A variety of different equilibria resulting from signaling in the international context can develop. It has been suggested that, over time, state practice might be separated into four rough categories: acceding and complying States, acceding and noncomplying States, nonacceding and complying States, and nonacceding and noncomplying States. Those States that accede and comply, of course, demonstrate the lowest discount rate, while those who neither accede nor comply show the highest likelihood to defect from cooperative relationships.<sup>133</sup>

Signaling theory, however, remains incompletely understood. In particular, it suffers from certain limitations on the ability to describe how States "agree" what behaviors are signals and about the content of these signals. A signal is only worthwhile if members of the community understand what it means. Advocates of signaling theory recognize this limitation and the particular problems of developing shared meaning in an international society that is itself comprised of many different cultures.<sup>134</sup> Should, for example, ratification be considered a signal? Should compliance? Should both? As a way of engaging this problem, it might be suggested that signaling is the result of "general perceptions of performance in an area rather than individual actions in individual cases."<sup>135</sup> But this begs the question of "whose perception": for if meaning has to be shared, then how does subjective perception matter? It further begs the question of which general perceptions amount to signals, and which do not.<sup>136</sup> Such difficulty in identifying the behaviors and content of signals results in

---

131. POSNER, *supra* note 46, at 17.

132. *Id.*

133. Moore, *supra* note 8, at 903.

134. *Id.* at 901-02 (internal citations omitted). While they cannot explain how the process happens, they rely generally on the fact that it is in nations' interests to agree on meaning so that signaling can take place.

135. *Id.* at 896.

136. See McAdams, *Signaling Discount Rates*, *supra* note 54, at 676 (arguing that the most efficient means of producing a reputation for a particular behavior is performing the behavior and that this reputational strategy will almost always overcome the signaling strategy, thus suggesting that few behaviors will be signals); see also Moore, *supra* note 8, at 896 (recognizing that if the costs of a signal are sufficiently low or the benefits sufficiently high then both states with high and low discount rates will signal).

a theory that is primarily descriptive. If a State's behaviors are perceived as cooperative then those behaviors are signals. If they are not perceived as cooperative then those behaviors are not signals.

Expressive law provides a more complete basis for understanding just how behaviors carry information regarding cooperativeness. This mechanism is somewhat different than the signaling mechanism. As previously described, expressive theory explains that States seek the esteem of others because of the operation of a need-reinforcement mechanism that results in an attraction to others with whom a State shares cooperative benefits. States are thus interested in what other States prefer and do not prefer, and norms simply reflect the aggregate preferences of group members. Norm-congruent behavior, in turn, signals ones' evaluation of group esteem. The more willing a State is to abide by less certain norms, the more it values group membership. Behavior does not thus carry information on discount rate but on the importance of group membership. A State that highly values esteem from the group is less likely to act in defiance of the group's norms.

In sum, expressive theory generally supports the notion that States signal their nature to other States through certain behaviors. Expressive theory, however, suggests that States ratify and comply with treaties not to provide information on discount rate, but to, instead, signal their attraction to the group. While evaluation of group esteem provides the basis for understanding a State's cooperativeness, certainty of preference will provide a basis for determining whether ratification or compliance will act as a signal. That is, the more a group of States identifies a preference for particular behaviors associated with a treaty, the more certainty a member of the group has that the behaviors will receive esteem.

### *E. Implications*

The expressive model provides a framework for considering both the decision to enter into a treaty regime, as well as the decision whether to comply with the regime afterwards. The model suggests that the pull of international society can exert significant influence on the willingness of States to enter into treaties. Moreover, while signals of compliance may be more ambiguous, and thus harder to understand, normative forces can also impact significantly the willingness to comply with these agreements.

Let us begin with a discussion of the influence exerted by international norms on ratification. The force is two-fold. First, a State must consider the esteem benefit it receives from ratification. Second, it must consider the foregone disapprobation it will receive by not

ratifying. Recall the example of a newly formed Islamic democracy discussed above, and assume that it has two relevant preferences. That State wishes to increase its esteem with others, and it also has a general desire to not recognize or advance the particular rights being advanced by the CEDAW as far as providing the franchise to women.<sup>137</sup> In such a case we would assume that the State would not ratify the treaty. Indeed, under this model we expect a State only to enter into the CEDAW if (as posited above) the normative benefit outweighs the sum of all other costs. While normative pressures can be very substantial, the general notion that norms on their own can result in ratification of a significant number of treaties is certainly thrown into question by such an understanding.

Now take into account the costs associated with not ratifying the CEDAW in terms of negative esteem. That is, by not joining the treaty regime, the newly democratic Islamic State will suffer esteem costs from failure to uphold the recognized international norm that favors women's equality while also gaining no other benefit.<sup>138</sup> Assume the amount of decreased esteem will roughly mirror the expected increase that would come from ratification. In such a case, in addition to foregoing the esteem benefit that would result from joining the treaty regime, that State undergoes esteem loss. Where this is true, the newly democratic Islamic State, despite its *ex ante* preferences, will be much more inclined to ratify the CEDAW. The picture painted by the expressive law model supports the general notion that normative influences can make rational States ratify treaties. Indeed, foregone reputational losses may be responsible for some heavy lifting in terms of creating incentives to join treaty regimes.

The theory further suggests that, to the extent the costs of making ambiguous the message sent by ratification are lower than the costs of complying with a treaty, a State will choose to do the former. Two major variables are relevant to the decision to ratify or comply for normative purposes: the effect of monitoring and the impact of

---

137. We could, of course, model specific beliefs regarding undertaking the behaviors required by a treaty. For example, a treaty standard requiring accommodating disabled individuals may carry with it a large number of underlying beliefs. Such a standard will cost money to the government in the form of accommodations and will slow the economy by taking business resources and the disabled are not capable workers compared with the rest of the workforce. The leader of that nation will evaluate each of these beliefs and its certainty to reach an ultimate belief about the behavior. For the sake of simplicity, we present this more complex calculus in terms of ultimate belief.

138. The negative reaction aimed at the United States for not supporting the Kyoto Protocol is another example of such disesteem. For an overview of the negotiation techniques and an evaluation of their likely success, see Jutta Brunnée, *A Fine Balance: Facilitation and Enforcement in the Design of a Compliance Regime for the Kyoto Protocol*, 13 TUL. ENVTL. L.J. 223 (2000).



resource limitations. An important variable and its effect on this key issue in current human rights scholarship is the effect of monitoring on decisions to enter treaties. Monitoring is the key to reputational enforcement. Under the expressive model, norms are endogenous; they reflect the preferences of other State members of a group. One follows such norms out of a desire for esteem (or to avoid disesteem). If a State could act contrary to the norm without surveillance, the State would receive no esteem or disesteem and thus would not be concerned regarding the normative consequences of its actions.<sup>139</sup>

The second major variable arises from the understanding that many States are not able to reach a treaty's goals due to financial or other limitations. As a result, many international treaties stipulate that States need only make progress that is reasonable in light of resources.<sup>140</sup> The existence of such a standard provides some basis for States to ambiguate the signal sent when human rights violations are found. They do so by recognizing that their behavior counters the norm and by expressing their concern, while also characterizing the behavior as previously unknown or fiscally unavoidable. In such a manner, the State maintains the status it receives through ratification and limits or completely neutralizes the disesteem it would otherwise receive for violating the norm. Because of the ability to make signals more ambiguous, one would expect States from whom progressive realization is required to feel less normative pressure to comply than to ratify.

Let us return to an earlier example of the five States deliberating whether to ratify a treaty providing immigrants with basic social entitlements to consider the interaction of these forces on a decision based on regime structure. At one end of the continuum is a regime without monitoring, but one that contains clear standards on measuring reasonable implementation. In such a case, ratification would likely be universal. This is because even the low esteem benefit that arises from agreeing to the terms of an essentially aspirational regime will outweigh the costs needed to implement it. However, as discussed previously, compliance will likely also be minimal. At the

---

139. This is similar to domestic examples of individuals who decry drinking, smoking, spousal abuse, or the use of pornography in public, but feel free to undertake such behaviors behind closed doors.

140. A seminal example is the International Covenant on Economic, Social and Cultural Rights, G.A. Res. 2200A, at 49, U.N. GAOR, 21st Sess., Supp. No. 16, U.N. Doc. A/6316 (Dec. 16, 1966). After setting out broad aspirations, the treaty mandates that "[e]ach State Party to the present Covenant undertakes to take steps . . . to the maximum of its available resources, with a view to achieving progressively the full realization of the rights recognized in the present Covenant." *Id.* at Article 2.

other extreme is a regime that requires complete compliance with its terms through strong monitoring procedures.

Assume for example that there are three different strategies available for providing immigrants with social rights. In ascending order of cost, these are the relatively low cost of requesting voluntary compliance with a rule empowering immigrants (inexpensive because it cannot be enforced), the moderate cost of providing health care and unemployment benefits to only those immigrants unable to sustain themselves (restrained, because most States will not allow these individuals entry in the first place), and the high cost of entitling immigrants to the same range of social benefits bestowed on all citizens (expensive due to the ratio of benefits provided to taxes paid). If a standard required implementation of the high cost strategy and failure to comply would almost certainly be discovered, the costs would be dramatically different. Certainty of esteem loss would be great and only those States for whom esteem benefits would outweigh the high costs of implementation would join the regime. For example, if the certain loss of esteem was less than compliance costs, that State would not join. One would expect such a regime to have high compliance but less universal ratification.

Reality, of course, occurs somewhere in between these extremes. Human rights treaties often have substantial formal monitoring requirements that are set off by limitations on their enforcement in the form of behavior that will satisfy the treaty's standards or exempt a State from compliance.<sup>141</sup> Assuming the goal of the process would be to gain as much compliance from as many States as possible, we would suggest a regime design that balances all of these concerns. As a threshold matter, monitoring in all cases must be strong because without monitoring one would expect minimal compliance. The most effective regime design would provide for specific, individually negotiated benchmarks for each country.<sup>142</sup>

---

141. For normative analysis, see Ryan Goodman, *Human Rights Treaties, Invalid Reservations, and State Consent*, 96 AM. J. INT'L L. 531 (2002); for examples, see citations *infra* note 151.

142. A variation on this first model would be the use of framework conventions that allowed nations to agree to the general desirability of a goal with negotiated standards to occur at a later time. Such a regime may allow for the certainty of the norm to be established and increased as more nations commit to and implement it—building increased pressure on the esteem function of other nations that are later to commit. One example of this type of convention is the European Union's directive prohibiting employment discrimination, the Council of Europe Directive Establishing a Framework for Equal Treatment in Employment and Occupation, Council Directive 2000/78, 2000 O.J. (L 303) 16 (EC). The European Union elected not to define "disability" on the ground that member States would define the term within their particular legal and cultural contexts. This in turn creates an incentive for a definition that is believed not to violate the umbrella directive. See Lisa Waddington, *Implementing the Disability Provisions of the Framework Employment Directive: Room for Exercising National Discretion*, in *DISABILITY*

Objectively verifiable standards, of course, would help increase the efficacy of monitoring. Such a regime would yield the largest number of ratifying States while also ensuring, to the extent possible, the highest standards of compliance for each country relevant to their desire for esteem.

To synthesize our analysis we provide a model of how the above insights on treaty design and monitoring may affect the implementation of the forthcoming United Nations Convention on the Rights of Persons with Disabilities. This treaty is expected to be adopted by the 61<sup>st</sup> General Assembly.<sup>143</sup> As of this writing, several provisions remain subject to negotiation, the most crucial of which focuses on the standards and mechanisms for implementing the treaty.<sup>144</sup>

To address the manner of the convention's implementation, a committee of National Human Rights Institutions ("NHRIs" are quasi-judicial bodies that implement human rights at the domestic level) submitted a draft proposal.<sup>145</sup> The NHRI submission contains innovative mechanisms directed at ensuring compliance. Most notable are the creation (in conjunction with disabled persons organizations) of baseline measurements and the drafting of individual national level action plans against which to assess progress, collective complaint mechanisms, and regional implementation.<sup>146</sup> Taken together, the proposed NHRI measures ensure transparent and efficient monitoring of the human rights of disabled persons. Although the adoption of all or part of this proposal is not guaranteed, its content is clearly illustrative of the reputational mechanism we model in this Article.<sup>147</sup>

---

RIGHTS IN EUROPE: FROM THEORY TO PRACTICE 107, 109 (Anna Lawson & Caroline Gooding eds., 2005) ("EC directives set goals which Member States are obliged to achieve within a given period, but leave Member States free to select the most appropriate means of achieving those goals.").

143. Information on the treaty is available at <http://www.un.org/esa/socdev/enable>. For an analysis by one of the participants, see Michael Ashley Stein, *Disability Human Rights*, 95 CAL. L. REV. (forthcoming 2007).

144. See Letter from the Chairman to all Members of the Committee (Oct. 7, 2005), available at <http://www.un.org/esa/socdev/enable/rights/ahcchairletter7oct.htm>.

145. Towards An Innovative Monitoring Mechanism for the Convention: Submission on Behalf of National Human Rights Institutions to the United Nations Ad Hoc Committee on a Comprehensive and Integral International Convention on the Protection and Promotion of the Rights and Dignity of Persons with Disabilities (August 2005) [hereinafter NHRI Submission]; available at <http://www.un.org/esa/socdev/enable/rights/documents/ihrscse.doc>.

146. *Id.* A conference was recently convened to discuss and refine the initial proposal. See [http://www.law.harvard.edu/news/2005/12/02\\_disabilities.php](http://www.law.harvard.edu/news/2005/12/02_disabilities.php).

147. Parenthetically, we note that current attempts at overhauling the United Nations treaty enforcement system are likewise harmonious with our framework. See The Secretary-General, *Strengthening of the United Nations: An Agenda for Further Change: Report of the Secretary General: Corrigendum*, U.N. Doc. A/57/387 (Oct. 16, 2002) (calling for measures aimed

The most effective way of using the pull of international society to leverage States into human rights compliance is through the creation of objectively-verifiable provisions that reflect realistic, realizable goals for each individual state member of the treaty regime.<sup>148</sup> The proposed NHRI model creates baseline measures and National Action Plans that are individual to every State party, and would be developed in conjunction with their administrative agencies, judiciaries, National Human Rights Institutions, and involved stakeholders (mainly, disabled persons organizations).<sup>149</sup> Consequently, the monitoring of a disability human rights convention would be transparent: a treaty body would measure a State's progress against the baselines and aspirations that a State had agreed-upon. Claims that a state could not realize such individually tailored goals will be less credible given the process by which the goals were created. Moreover, the use of regional assessments provides added incentive for a given State to comply with the convention to at least a level equal to that of its geo-political neighbors.<sup>150</sup> In sum, this form of monitoring puts into place documents that enable the questioning and enforcement of a State's commitment by specifically delineating the progress that must be achieved to maintain that State's reputational benefit. Even if that State argues that economic circumstances have changed since ratification such that compliance has been made more difficult, this system still constitutes a vast improvement over the laudable but vague aspirations contained in current treaties.

Having provided a model of how expressive forces can best be used for achieving ratification and compliance, let us now turn to one or two further implications of the theory before providing our example of how the theory will be operationalized. One of the major implications of expressive theory for international law hinges on the important notion that group identity and attraction matters. There are at least two observations that flow from this understanding. First, for human rights treaties to succeed fully, States must be involved in other meaningful cooperative relationships. Without trade and other interaction, States will not develop the attraction to the group that

---

at increasing coordination among monitoring bodies, greater standardization of reporting requirements, and additional monitoring at the national level).

148. This is a commonly held assertion. See, e.g., George H. Aldrich, *Compliance with the Law: Problems and Prospects*, in EFFECTING COMPLIANCE 162 (Hazel Fox & Michael Meyer eds. 2000) (transparent monitoring mechanisms necessary to effectuate regime compliance); Christopher C. Joyner, *Sanctions, Compliance, and International Law: Reflections on the United Nations' Experience Against Iraq*, 32 VA. J. INT'L L. 1, 37-46 (1991) (positing that only an effective enforcement mechanism can alter regime behavior).

149. See NHRI Submission, *supra* note 145.

150. *Id.*

results from the operation of the need-reinforcement principle. In this sense, strategies to protect human rights must be drawn much more broadly than simply within the context of human rights institutions. Aggressive pursuit of cooperation with States must be considered a part of any effective human rights protection scheme, at least to the extent such a scheme relies on normative forces for its success.

Second, failure to identify with a group can lead States to esteem behaviors that are contrary to international norms. The rebel phenomenon is well known. Rebels identify not with the main group but with a group of "outsiders;" indeed, rebels in some measure define their identities by the amount they contravene established norms. Accordingly, rebels smoke, drive quickly and recklessly, or ride their motorcycles without helmets to flaunt the norms of the majority. States that act as rebels get esteem from a group other than the majority and thus get utility from not following a norm. International lawmakers must be wary of this phenomenon and factor it into their regime design. Creating broad, well-designed regimes will positively impact in-group members but will spur decreased compliance from self-defined rebels. In its broadest sense, this idea parallels the argument above that the most effective course for compliance is building group allegiance, but it is important to recognize that regimes are not universal. Thus, a true measure of a treaty's effectiveness will not simply be the compliance it engenders within the group but also the decreased compliance of those outside the regime.

#### *F. Making Human Rights Treaties Matter*

In an influential article, Oona Hathaway analyzes substantial empirical data and concludes that, in certain cases, States entering human rights treaties actually increase their violations of treaty standards after joining.<sup>151</sup> Hathaway's conclusion raises concerns about the efficacy of human rights treaties specifically, and also throws into relief general concerns regarding the ability of social forces to influence compliance.

Hathaway provides an explanation for these results that is grounded in expressive theory. She suggests that States behave in this manner in order to receive the reputational benefit that comes from joining treaty regimes without having to bear related compliance costs. Both Hathaway's empirical methodology and her theoretical model have been criticized.<sup>152</sup> Some commentators, for example, take issue with Hathaway's reliance on observed abuses. They note that,

---

151. Hathaway, *Do Human Rights Treaties*, *supra* note 5, at 1940.

152. See citations *supra* note 6.

because monitoring increases once a treaty has been ratified, the number of observed treaty violations will also increase, in which case ratification correlates with greater knowledge of existing abuses rather than with greater abuses.<sup>153</sup> A core assumption of her theory—that failure to comply does not carry reputational effects—also raises expressive law concerns. Hathaway's critics argue that reputational benefits decrease as a State is revealed to be out of compliance with treaty terms. Thus, they point out that, a claim that reputational factors can create incentives to enter into a human rights treaty without also intending to comply with its mandates, is misplaced.<sup>154</sup>

The lack of a full theory of expressive international law, however, has rendered impracticable further discussion of this reputational explanation. Without a comprehensive theory one cannot begin to consider whether Hathaway's empirical findings are simply incorrect or, if they are correct, whether they reflect poor treaty design or the inability of social forces to influence the behavior of States. Our expressive theory picks up this challenge and provides insights into Hathaway's arguments. It agrees that, if monitoring is as low or non-existent as she claims, many States motivated by reputation will ratify treaties without complying afterwards. It also provides an understanding of when the backsliding Hathaway identifies may occur. Ultimately, the expressive theory provides reasons for questioning Hathaway's empiricism. It also clearly suggests that, to the extent existing treaties are not effective, this shortcoming is the result of their ineffective design, and not proof that social forces cannot influence State behavior.

Hathaway's expressive explanation of her data can be described by syllogism. First, she argues, States ratify treaties for expressive purposes that function like a roll call vote in Congress. As she explains it, ratification is a "pleasing statement not necessarily intended to have any real effect on outcomes. It declares to the world that the principles outlined in the treaty are consistent with the ratifying government's commitment to human rights."<sup>155</sup> Second, she argues that when monitoring and enforcement are minimal, a disjuncture may arise between the expressive and instrumental roles of treaties.<sup>156</sup> Hathaway concludes that in instances of disjuncture,

---

153. Goodman & Jinks, *Measuring the Effects*, *supra* note 6, at 174-76 (stating that Hathaway's model cannot distinguish between a State in which levels of torture increase post ratification and a State in which torture declines post ratification but appears to increase because liberalization eases the process of documenting and reporting instances of torture).

154. *Id.* at 179-80 (describing the paradox that to be meaningful the signal must be costly yet at the same time States will choose to ratify if the ratification signal is relatively costless).

155. Hathaway, *Do Human Rights Treaties*, *supra* note 5, at 2005-06.

156. *Id.* at 2006.

treaties' expressive aspects can "relieve pressure for real change in performance in countries that ratify the treaty," because these instruments "offer rewards for positions rather than for effects."<sup>157</sup> A corollary to this main argument flows from her observation that regional treaties are more likely to create conditions for backsliding than universal ones. She concludes that this dynamic results from even greater pressure for States to join treaties when those treaties concern other States with whom they share strong regional ties.<sup>158</sup>

Our theory suggests that Hathaway's analysis may be right in very specific situations where States are certain that failure to comply will be of no consequence, and a strong international norm existed before a treaty regime developed. Consider, for example, a State before the creation of a treaty limiting torture of prisoners. Assume that before the treaty, a well-known norm exists against this practice. Concerns for esteem would lead the State to signal by behaving a certain way—in this case, not torturing prisoners. Treaty ratification, however, may serve to replace actual behavior as a way of signaling commitment to international society. When that circumstance occurs, the need to act to signal is decreased, and one would expect a State to adjust its behavior accordingly.

Hathaway's theory works, however, only in cases where a State can act in derogation of a treaty without serious consequence to its reputation. For example, in situations where the State can plausibly argue that it is progressively addressing the treaty's stated goals or where its acts will not be discovered. In these circumstances, the benefits of esteem received from ratification must be balanced against the loss that comes from compliance failure. Where the net loss between these two is greater than the esteem received before entering into the treaty, the State will backslide. Note, however, how quickly the reputational effects assert themselves when monitoring cannot be discounted or where the standards in the treaty create objective and empirically verifiable standards of conduct through domestic action plans. In that case, failure to abide by the treaty not only carries reputational impacts because the State cannot as easily ambiguate the message carried by its behavior, but now the State will also be seen as violating another (more global) norm—the norm of respecting and keeping one's obligations. In such cases, it is highly unlikely that backsliding would occur.

Expressive theory thus explains that backsliding may occur in certain cases. We take issue with Hathaway, however, over how prevalent such situations are, and we disagree with her to the extent

---

157. *Id.* at 2007 (internal citations omitted).

158. *Id.* at 2016.

that she claims failure to comply with most treaties would have little or no reputational effect. In particular, we agree with Hathaway that self-monitoring by States of their own compliance with human rights treaties leaves much to be desired.<sup>159</sup> We disagree, however, with her exclusive focus on “woefully inadequate” State-driven compliance mechanisms when examining the crucial issue of monitoring. In pointing out this inadequacy, Hathaway concludes that “the failure of a country to comply with its treaty obligations is, in most cases, unlikely to be revealed and examined except by already overtaxed NGOs.”<sup>160</sup> Yet it is precisely NGOs, even when over-taxed, that exist to identify and report human rights violations. Moreover, the dissemination of this information, and the reputational affects that it conveys, leverage States to alter their behavior in order to avoid losing the esteem of the global community.<sup>161</sup>

Our expressive theory provides robust support for this understanding. Every behavior provides a signal to the extent a State will infer the group commitment of the signaling State from it. While signals of group commitment communicated through compliance can be ambiguated, the notion that failure to comply carries with it little or no reputational effect is seriously circumspect. We would, in particular, take issue with claims that all or most States can plausibly deny their commitment to treaties over a period of time. This is especially true in relation to the previously discussed notion of progressive realization. As States develop economically over time, the ability to explain away a consistent level of non-compliance becomes more difficult. One would thus expect that developing nations would decrease their violations over time, and that developed nations would be unable to explain away many violations. In consequence, further inquiry needs to be made both into Hathaway’s empirical methodology, and into its normative implications.

---

159. Hathaway argues that this type of monitoring is “woefully inadequate.” *Id.* at 2008. See also PETER UVIN, HUMAN RIGHTS AND DEVELOPMENT 140 (2004) (finding that monitoring processes “constitute some of the most powerless, under-funded, formulaic, and politically manipulated institutions of the United Nations”). We note, parenthetically, that the current United Nations treaty monitoring system is undergoing reform.

160. Hathaway, *Do Human Rights Treaties*, *supra* note 5, at 2008 (internal citations omitted).

161. Moreover, the United States State Department also compiles information on national human rights that it publishes annually. An electronic archive of Country Reports on Human Rights Practices compiled prior to 2001 is available at [http://www.state.gov/www/global/human\\_rights/hrp\\_reports\\_mainhp.html](http://www.state.gov/www/global/human_rights/hrp_reports_mainhp.html). The irony, then, of Hathaway’s claims is that although she describes the difficulty of gathering information on human rights compliance she herself has gathered the information in a manner that allows her to measure compliance in a detailed and complex manner.



## CONCLUSION

This Article provides an initial and comprehensive understanding of the forces behind human rights treaty creation and compliance by drawing on evolving expressive law literature. It began by setting forth a model that explained how normative pressure influences rational actors to alter their behavior and beliefs while seeking regard from other group members. Next, the Article applied the model to State treaty ratification and compliance, and argued that the powerful pull of international society creates a rational choice for States to abide by these instruments because of their desire to be part of, and esteemed by, the global community. It demonstrated how international process influenced States' preferences and normative beliefs, which in turn affected their willingness to ratify and comply with treaties. The Article also demonstrated how expressive theory harmonizes the valuable (but incomplete) contributions of divergent international law scholars into a more comprehensive theory of why States enter into and honor international agreements. In doing so, it presented a detailed framework from which predictable determinations can be drawn. Finally, the expressive framework was used to describe the means for harnessing the "pull of international society" in the specific context of a disability human rights treaty that one of the authors is drafting and which will soon be presented to the United Nations General Assembly.

Identifying how the desire for esteem from the global community motivates States to comply with international law has eluded scholars since the time of Grotius. The current debate among international legal scholars focuses on the efficacy of a reputational mechanism to impel State compliance with human rights treaties. The resulting literature is an important body of work that recognizes the centrality and implications of esteem-based mechanisms on compliance, but fails to provide a level of detail necessary to understand or predict how the desire for esteem affects the behavior of States. This Article has addressed that lacuna by presenting a comprehensive and detailed expressive law framework. In doing so, it provided an avenue for continued (and more focused) discussion and debate over how the desire for esteem influences States to enter into and comply with human rights treaties.

It is imperative that we understand the impact that the pull of international society exerts over State behavior. The United Nations is currently undergoing a systematic reassessment, with special emphasis on the structure and efficacy of its treaty system. Dramatic changes are being proposed, including the combination of all existing treaty commissions (including the proposed disability-related one) into

a single super-treaty monitoring body. The final (structural) configuration of the United Nations treaty system, however, is less significant for influencing human rights compliance and regime change than is the efficacy of its monitoring processes. Transparent and efficient implementation of individually crafted and objectively verifiable standards must be created if States' need for international esteem is to be leveraged towards human rights treaty ratification and compliance. Expressive principles supply the means to harness the pull of international society and make treaty ratification and compliance effective.

\*\*\*