

Spring 2015

Norms, Attitudes and Compliance

Stephen R. Galoob
University of Tulsa College of Law

Adam Hill

Follow this and additional works at: <http://digitalcommons.law.utulsa.edu/tlr>

 Part of the [Law Commons](#)

Recommended Citation

Stephen R. Galoob, & Adam Hill, *Norms, Attitudes and Compliance*, 50 Tulsa L. Rev. 613 (2015).

Available at: <http://digitalcommons.law.utulsa.edu/tlr/vol50/iss2/8>

This Book Review is brought to you for free and open access by TU Law Digital Commons. It has been accepted for inclusion in Tulsa Law Review by an authorized editor of TU Law Digital Commons. For more information, please contact daniel-bell@utulsa.edu.

NORMS, ATTITUDES, AND COMPLIANCE

Stephen R. Galoob and Adam Hill*

GEOFFREY BRENNAN, LINA ERIKSSON, ROBERT E. GOODIN, & NICHOLAS SOUTHWOOD, *EXPLAINING NORMS* (2013). Pp. 290. Hardcover \$55.00.

INTRODUCTION

The connection between law and norms has been one of the most widely discussed issues in legal scholarship since at least the publication of Robert Ellickson's seminal work on the subject.¹ This topic has inspired (and, to an extent, unified) doctrinal scholars, legal theorists, economists, and sociologists.

Yet despite the breadth of these discussions, legal scholars of norms have left many fundamental questions unresolved. What are norms? How do they affect what people do? Are norms mainly social phenomena, or do they include legal phenomena as well?

Explaining Norms is a work of philosophy of social science by Geoffrey Brennan, Lina Eriksson, Robert Goodin, and Nicholas Southwood.² (For the sake of convenience, we refer to the authors collectively as "BEGS.") The book is a breakthrough, containing many important advances in thinking about norms. We use this review of *Explaining Norms* as an occasion to question the reductive, behaviorist definition of norms that predominates among legal scholars and influences much empirical legal research.

Part I of this review discusses two foundational questions about norms: what are they, and what is it to comply with them? BEGS provide a powerful account of the nature of norms and offer several insights about compliance.

Part II sets out the leading definition of norms in the legal literature, which we call "reductive behaviorism." Reductive behaviorism sees norms primarily as patterns or regularities of behavior. An implication of this definition is that to behave in accordance with a norm is to comply with it.

Part III deploys BEGS's insights about compliance to identify two important phenomena that reductive behaviorism leaves out. First, reductive behaviorism denies that norms could concern anything other than externally manifest behaviors, such as patterns

* Stephen Galoob is Assistant Professor of Law at the University of Tulsa College of Law. Adam Hill is a Ph.D. candidate in the School of Law at the University of California at Berkeley. The authors wish to thank Sam Halabi, Matt Lamkin, Ethan Leib, Richard McAdams, Brian Sheppard, and Bob Spoo, and the editors of the *Tulsa Law Review* for their attention. Special thanks to David DiPianto, whose comments inspired revision of substantial portions of parts II and III. Galoob's work was supported in part by a Dean's summer research grant from the University of Tulsa College of Law.

1. Robert C. Ellickson, *Of Coase and Cattle: Dispute Resolution Among Neighbors in Shasta County*, STAN. L. REV. 623 (1986); ROBERT C. ELICKSON, *ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES* (1991).

2. GEOFFREY BRENNAN ET AL., *EXPLAINING NORMS* (2013) [hereinafter *NORMS*].

of deliberation. However, many norms *do* concern deliberation, so reductive behaviorism cannot describe how these norms operate. Second, complying with certain norms requires not only that someone behave and deliberate in accordance with the norm, but also that the norms (or some higher-order commitment involving the norms) factor into the agent's deliberations. In these cases, to use BEGS's terminology, compliance requires *following* a norm, rather than merely *conforming* to it. Yet, reductive behaviorism cannot appreciate any meaningful difference between conforming to and following a norm, let alone explain why some norms must be followed. Thus, reductive behaviorism provides an incomplete account of how norms work. BEGS's insights can help remedy this conceptual deficiency and inspire improvements in the empirical study of norms.³

I. TWO BIG QUESTIONS ABOUT NORMS

Explaining Norms has three main parts. Part I analyzes the nature of norms. Part II addresses how norms emerge, persist, and change over time. Part III concerns compliance with norms. We elaborate (and criticize) the arguments in parts I and III of the book, which seem to be its farthest-reaching contributions.

A. *What are norms?*

1. Background

BEGS articulate and defend an account that they call the "norms as normative attitudes" view.⁴ Before discussing this view, two background observations are helpful to explain the ambitions of BEGS's project.

First, BEGS observe that norms have both a "normative" element and a "socio-empirical" element.⁵ The normative element is that norms implicate and are made up of "normative principles," which are presented as "general requirements" for those over whom the norm applies.⁶ This normativity distinguishes norms from statistical generalizations, which do not necessarily have any normative content.⁷ The socio-empirical element of norms means that they operate over and are "somehow accepted within" particular groups or communities.⁸ Whether a particular normative principle is actually the norm of a group is a "social fact" that is (at least in principle) subject to verification. This facticity depends on whether the principle has been accepted, not whether it is objectively valid. An objectively invalid normative principle (such as the principle that young girls should not be educated) could nevertheless be the norm of a group if it were accepted within the group. For BEGS, a satisfactory definition of norms must explain both these normative and socio-

3. To be sure, the significance of this shortcoming for scholars will vary across types of empirical research. Some empirical research (e.g., purely data-driven projects) does not posit or rely on any definition of norms at all. The soundness of an underlying definition of norms might be immaterial for those projects.

4. NORMS, *supra* note 2, at 15.

5. *Id.* at 3.

6. *Id.*

7. *Id.* at 2.

8. *Id.* at 3.

empirical elements.⁹

Second, BEGS analyze norms through the lens of rational choice theory, which is characteristic of much economic analysis. Rational choice theory assumes “that agents are preference satisfaction maximizers” and that the effects of norms are explained by “the content of those preferences.”¹⁰ This approach contrasts with some sociological approaches, which give pride of place to structural considerations in describing what norms are and how they operate. BEGS also adopt a specific picture about why people act that is associated with the eighteenth century Scottish philosopher David Hume.¹¹ On this Humean picture of motivation, people act in order to satisfy their desires.¹²

2. The “Norms as Normative Attitudes” View

On the norms as normative attitudes view, norms are “clusters of normative attitudes,” in that “[w]hat it takes for a . . . normative principle to be accepted in and hence to be the norm of a group is, at least in part, that a significant proportion of the members of that group have certain attitudes,” that these normative attitudes track (and are prompted by) the content of the norm, and that these attitudes meet certain standards of publicity.¹³ More formally, a normative principle, P, is a norm within a group, G, if and only if:

- (i) a significant portion of the members of G have P-corresponding normative attitudes; and
- (ii) a significant proportion of the members of G know that a significant proportion of the members of G have P-corresponding normative attitudes.¹⁴

This definition involves several technical terms. “Normative attitudes” are a certain kind of principled judgment about the world. For BEGS, normative attitudes include judgments about the normative status of actions (such as whether they are forbidden or required); normative expectations concerning the actions of others; reactive attitudes (such as resentment and indignation) and “dispositions to have such attitudes”; and attitudes that presuppose normative judgments, normative expectations, and reactive attitudes.¹⁵ A normative attitude “corresponds” to a principle when it “appropriately reflects the content and normative force” of that principle.¹⁶ BEGS leave open what counts as a “significant proportion” of a population, and they allow that significance might vary for different types of norms.¹⁷ Nor do BEGS specify what it means for someone to “have” a normative attitude,

9. *Id.* at 3-4.

10. *Id.* at 8.

11. *Id.* at 9, 33, 198-99.

12. *Id.* at 198-99 (citing MICHAEL SMITH, *THE MORAL PROBLEM* (1994)). The Humean picture of motivation also asserts that the reasons a person has or believes herself to have are “motivationally inert” and do not explain why she behaves in the ways that she does. *Id.* at 198.

13. *Id.* at 29.

14. *Id.* at 31.

15. *Id.* at 29.

16. *Id.*

17. *Id.* at 30.

although they deny that having a normative attitude necessarily prompts one to act in a particular way.¹⁸

To illustrate the norms as normative attitude view, consider a normative principle like “ordinarily keep your promises.” This principle has been accepted in some communities, and so is a norm for those groups. If someone violates the principle (say, by breaking a promise solely in order to obtain a material gain), then she is judged by a significant proportion of the population to have acted wrongly and it is appropriate for the promisee to resent her violation. If these normative attitudes do not attach to a principle in a group, then the principle is not a norm for that group. In other words, “keep your promises” is not a norm in communities where breaking a promise does not occasion these kinds of critical judgments and attitudes.

3. Four Aspects of the “Norms as Normative Attitudes” View

Here are four aspects of the norms as normative attitudes view that contrast with the predominant definition of norms in the legal literature. First, the norms as normative attitudes view is non-reductive because it invokes essentially normative concepts to explain why a normative principle is a norm. By contrast, BEGS contend, reductive approaches hold that whether a principle is a norm for a group turns on social facts that do not essentially involve normative concepts.¹⁹

Second, the norms as normative attitudes view treats normative attitudes as bedrock.²⁰ If normative attitudes do not attach (in the right way, among the right people in a

18. *Id.* at 38.

19. *Id.* at 28-29.

20. BEGS contrast the norms as normative attitudes view with two reductive approaches, which they call the “norms as practices” and “norms as desires” views. *Id.* at 15.

On the “norms as practices” view, a normative principle is a norm when a significant proportion of members behave in a way that is consistent with the principle; this behavior is explained, in part, by group members having positive attitudes toward behavior that conforms with the principle; and these attitudes are a matter of common knowledge. *Id.* at 16. BEGS contend that, while norms and social practices typically go together, the latter are neither necessary nor sufficient to explain the existence of the former. To show why practices are unnecessary, BEGS offer an example to illustrate the possibility that a norm could exist (in their example, a norm against urinating in public swimming pools), even though there is no widespread social practice of people behaving in a way that conforms to it (that is, where people routinely urinate in public swimming pools). *Id.* at 20. BEGS argue that the norms as normative attitudes view can explain why there is a norm in this example: namely, because members have normative attitudes that correspond to the norm (*e.g.*, the judgment that “one mustn’t urinate in public swimming pools” and a disposition to “disapprove of those who do”), even though the behavior of most members does not correspond to the norm. *Id.* at 31.

On the “norms as desires” view, norms are clusters of desires and (on some versions) beliefs about the desires of others. *Id.* at 22. BEGS focus Cristina Bicchieri’s account, which sees norms as clusters of “desire[s] to comply with [] normative principle[s],” conditional on others also complying with them and expecting compliance. *Id.* at 23-24 (citing CRISTINA BICCHIERI, *THE GRAMMAR OF SOCIETY: THE NATURE AND DYNAMICS OF SOCIAL NORMS* (2006)). Here, too, BEGS provide examples to suggest that clusters of desires are neither necessary nor sufficient for the existence of a norm. In the case of “the Chastians,” virtually all the members of a group behave in a chaste manner, yet have a conditional desire to live under a norm that mandates unchaste behavior (should others behave unchastely and expect general behavior in this way). *Id.* at 25. BEGS contend that the Chastians do not have a norm of unchaste behavior, even though Bicchieri’s account would suggest that they do. *Id.* Likewise, BEGS invoke the case of “the Philanthropians” (who have a norm requiring donation of income, even though virtually all the members of the group prefer to spend their income on luxury goods) to argue that clusters of desires are insufficient to show the existence of a norm. *Id.* at 27-28. By contrast, the norms as normative attitudes view can explain why the Chastians do not have a norm of unchaste behavior (because the Chastians do not themselves have normative attitudes that correspond to the principle that one should behave unchastely), as well as why the Philanthropians have a norm regarding philanthropy (because they disapprove of

group) to a principle, then that principle is not a norm for that group.

Third, the “core function” of norms is to facilitate accountability: norms allow group members to possess a “recognized right or entitlement to determine how one is to behave.”²¹ This position differs from the predominant view that norms primarily serve to facilitate coordination (by allowing groups to “reach[] mutually beneficial outcomes where [their] interests are fairly well aligned but where there are multiple different ways in which [they] could achieve such outcomes”) and/or cooperation (by allowing groups to reach “mutually beneficial outcomes” in situations where the interests of members “are pitted against one another, but not intractably so”).²² BEGS defend their position by arguing that norms can be important even where facilitating cooperation or coordination is “either impossible or unimportant.”²³

Fourth, the norms as normative attitudes view describes norms generally. BEGS distinguish between “formal” and “non-formal” norms.²⁴ Legal norms fall into (but do not exhaust) the former camp; moral and social norms fall into the latter.²⁵ BEGS contend that, while formal and non-formal norms can overlap, their operation differs across these contexts. Formal norms are accompanied by secondary rules that govern interpretation and enforcement, while non-formal norms lack these secondary rules. In other words, formal norms have an identifiable source, and that source (typically) has the authority to interpret and enforce the norm. Formal norms therefore create accountability relationships that are mediated through an external authority.²⁶ The normative attitudes that attach to formal norms can concern either the act picked out under the normative principle, or the authority of an entity to interpret or enforce the norm, or both.²⁷ By contrast, non-formal norms do not typically have a definitive origin or an authoritative interpreter. Rather, interpretation and enforcement are decentralized.²⁸ Non-formal norms therefore generate an unmediated form of accountability because members of a group enforce them against each other. Among non-formal norms, BEGS distinguish moral norms from social norms, which differ based on the grounds on which the normative principles and attitudes are justified. Moral norms are grounded on judgments that are independent of social practices, while social norms are grounded (at least in part) by practice-dependent social attitudes.²⁹

In sum, four important features of the norms as normative attitudes view are that it is non-reductive (because it uses normative concepts to explain the normativity of norms); it cashes out norms in terms of the normative attitudes related to normative principles; it sees the core function of norms as fostering accountability; and it is capable of describing legal, moral, and social norms.

the non-philanthropic behavior of others, despite harboring non-philanthropic desires themselves). *Id.* at 31-32.

21. *Id.* at 35-36.

22. *Id.* at 35.

23. *Id.* at 36.

24. *Id.* at 37.

25. *Id.* at 40-41.

26. *Id.* at 52.

27. *Id.* at 46-47.

28. *Id.* at 53-55.

29. *Id.* at 58-59, 67-81.

B. *Complying with norms*

In Part III of *Explaining Norms*, BEGS address issues related to complying with norms.

1. What Do Norms Govern?

Norms typically govern behavior. However, BEGS argue, norms can also “bear[] upon what goes on inside people’s heads”³⁰ by “demand[ing] that we have or form certain attitudes and that we think or deliberate in certain ways.”³¹ In other words, BEGS claim that at least some norms are what we will call *deliberation-sensitive*; whether one lives up to these norms turns, at least in part, on how one deliberates.

BEGS describe several ways that norms can be deliberation-sensitive. For example, if (as Joseph Raz contends) legal norms purport to provide exclusionary reasons, then they might be taken to demand that adherents exclude certain kinds of considerations from deliberation, to “think no further of them.”³² A norm could also govern deliberation by removing some alternatives from the agenda of available options³³ or “narrow[ing] the agenda down to a small number of alternatives.”³⁴ Some norms forbid (and others require) “calculative reasoning,” or scrupulous thinking about the costs or benefits of a particular kind of decision.³⁵

BEGS contend that non-formal norms can impose deliberative requirements, while formal norms do not.³⁶ According to BEGS, formal norms typically apply only to actions, and “there would be something very odd in the case of laws that make demands on our attitudes.”³⁷ Although BEGS do not address this question directly, they also seem to contend that formal norms are deliberation-insensitive. It is unclear whether BEGS see deliberation-insensitivity as a conceptual feature of formal norms or a contingent feature of Western legal systems. (BEGS contemplate the deliberation-insensitivity of formal norms only briefly in Chapter 3, and this topic is not mentioned in their more sustained discussion of norms and deliberation in Chapter 12.) By contrast, BEGS argue, moral and social norms seem capable of “involv[ing] a much richer set of objects,” including the reasons for which someone acts.³⁸

BEGS’s analysis of the deliberation-sensitivity of norms is intriguing, but incomplete and (in places) incorrect. It is incomplete because BEGS focus primarily on how norms might prohibit certain forms of deliberation. They do not extensively discuss the possibility that a norm could impose more direct requirements on how someone deliberates—for example, by mandating that she attend (or attribute practical significance) to

30. *Id.* at 193.

31. *Id.* at 245.

32. *Id.* at 250-51. This gloss on Raz’s position is inaccurate, since Raz denies that exclusionary reasons are “reason[s] to avoid thinking, considering or attending to certain matters,” or that they preclude an agent from engaging with excluded first-order considerations. See JOSEPH RAZ, PRACTICAL REASON AND NORMS 184-85 (1999).

33. NORMS, *supra* note 2, at 251.

34. *Id.* at 253.

35. *Id.* at 254-67.

36. *Id.* at 50.

37. *Id.* BEGS do not further elaborate why it would be odd for laws to make demands on deliberation.

38. *Id.* at 50-51.

certain kinds of considerations. Nor do BEGS explain when or why any norm might be deliberation-sensitive. It is incorrect because some formal norms are deliberation-sensitive. As noted below,³⁹ a variety of legal norms (including, notably, fiduciary duties) govern both behavior and deliberation.

2. Modes of Compliance

BEGS distinguish complying with a norm from breaching a norm (that is, behaving or deliberating contrary to its requirements).⁴⁰ They also identify two different modes of complying with a norm: “following” and “conforming.” Someone follows a norm when she not only behaves (or deliberates) as the norm requires, but also justifies these actions by the fact that the norm requires them.⁴¹ Someone conforms to a norm when she behaves or deliberates as the norm requires, “not because of the norm, but because of other considerations associated with the norm.”⁴² Let’s call the difference between these two modes of compliance the *following/conforming distinction*.⁴³

As the definition of these terms suggests, following is a more demanding modality than conforming. According to BEGS, following a norm requires “internalizing” it, or becoming disposed to “treat the norm as a non-instrumental reason” for action.⁴⁴ For example, take the norm to ordinarily keep one’s promise. If Ann behaves and deliberates exactly as this norm requires, but does so for reasons that do not essentially reference the norm (say, because her deity requires it), then she would not follow the norm of promise keeping. By contrast, to conform to a norm is to “externalize” it, or to treat it as providing a “purely instrumental reason to act in accordance.”⁴⁵ Ann’s same patterns of deity-inspired behavior and deliberation would count as conforming to the norm of promise keeping. Therefore, whether Ann complies with the norm of promise keeping depends on whether this norm must be followed, or if it can be satisfied through conformity.

After drawing the following/conforming distinction, BEGS utilize a Humean picture of norm internalization⁴⁶ to draw a number of stark conclusions. First, they contend that following formal norms is rare because internalizing legal norms is rare. People generally lack non-instrumental desires to act in accordance with legal norms and, in any event, these desires seldom explain why someone acts.⁴⁷ BEGS also contend that following legal norms is irrational. Even if a considerable portion of citizens take the state to be justified in enforcing a law, it does not follow that anyone should have a non-instrumental desire to act in accordance with that law.⁴⁸ Thus, people paradigmatically comply with legal norms by

39. See *infra* notes 98-109 and accompanying text.

40. NORMS, *supra* note 2, at 234.

41. *Id.* at 195.

42. *Id.* at 218.

43. This distinction parallels a similar distinction that Joseph Raz draws between complying with a norm and conforming to it. See RAZ, *supra* note 32, at 178; Scott Hershovitz, *Legitimacy, Democracy, and Razian Authority*, 9 LEGAL THEORY 201, 202 (2003); DAVID OWENS, SHAPING THE NORMATIVE LANDSCAPE 15 (2012); Stephen R. Galoob & Ethan J. Leib, *Intentions, Compliance, and Fiduciary Obligations*, 20 LEGAL THEORY 106, 111 (2014).

44. NORMS, *supra* note 2, at 219.

45. *Id.*

46. See *supra* notes 11-12 and accompanying text and *infra* notes 57-61 and accompanying text.

47. NORMS, *supra* note 2, at 210-12.

48. *Id.* at 211-13.

conforming to them,⁴⁹ and this mode of response is “perfectly rational.”⁵⁰

Second, BEGS see following a moral norm as similarly rare and irrational. BEGS invite us to consider the moral norm against rape.⁵¹ To follow this norm is to refrain from rape and to justify this refraining because it is prohibited by the norm (rather than for other moral considerations, such as that rape degrades and harms the victim). Yet, on the norms as normative attitudes view, a principle is a moral norm in a society only if “a significant proportion of the members of the society possess norm-independent normative attitudes corresponding to the rule or principle.”⁵² If a substantial portion of a society followed a moral norm, then it would cease to be a moral norm because people would treat the norm itself (rather than the norm-independent normative attitudes that correspond to it) as justifying the action.⁵³ Therefore, BEGS conclude, conformity is the “paradigmatic mode” of complying with moral norms.⁵⁴

Third, BEGS conclude that some social norms are capable of being followed. Because social norms are non-formal, they implicate normative attitudes about the justificatory force of the principles themselves, rather than attitudes about the justification of an entity’s authority to interpret or enforce the principles. Since the attitudes undergirding social norms are essentially practice-dependent, there is nothing irrational or self-defeating about following a social norm.⁵⁵ However, BEGS argue, nearly every norm that is satisfied by following can also be satisfied by conformity: to judge that a principle is a social norm is to also take oneself to have instrumental reasons (*e.g.*, reasons to avoid the disapproval of others) to act in accordance with it.⁵⁶

BEGS’s conclusions about the following/conforming distinction are questionable, and their arguments for them are problematic. Their case for the irrationality of following formal or moral norms stems from their Humean picture of internalization, on which someone internalizes a norm only if she develops “a non-instrumental desire to act in accordance with [the] norm”—that is, a desire in which “part of the justification” for acting in accordance with a norm is “simply the norm itself.”⁵⁷ The notion of a “non-instrumental desire to act in accordance with a norm” seems both mysterious and *ad hoc*. BEGS do not establish what such a desire is, whether non-instrumentality is a feature of a desire or part of its content, how non-instrumental desires differ from other conative states that are not required for internalization (such as “basic” desires),⁵⁸ or how one develops or comes to have a non-instrumental desire. Yet BEGS treat the mysteriousness of non-instrumental desires not as an infelicity in their argument, but rather as support for their conclusion. In

49. *Id.* at 213.

50. *Id.* at 229.

51. *Id.* at 214.

52. *Id.* at 215.

53. *Id.* at 230.

54. *Id.* at 215.

55. *Id.*

56. *Id.* at 232.

57. *Id.* at 201-02.

58. *Id.* at 201. BEGS suggest that the difference between non-instrumental and basic desires is that a non-instrumental desire can appear to be basic, but might not in fact be. *Id.* at 202. Yet this argument appears to be at odds with what BEGS call the “grounds view” concerning the difference between moral and social norms, *id.* at 58-59, since the grounds for non-instrumental and basic desires appear to be the same—namely, the agent’s belief that the desire is basic.

essence, they argue that it is implausible that anyone could develop non-instrumental desires regarding a moral or legal norm, and so it would be irrational for anyone to ever follow such norms (and, by implication, implausible for any norm to call for following as its paradigmatic mode of compliance).

There are several difficulties with this argument. First, BEGS never defend the Humean picture of internalizing a norm. Rather, they help themselves to it based on their (also undefended) choice to analyze norms through a rational choice framework.⁵⁹ This dialectical move is problematic because, as BEGS acknowledge, the Humean picture of motivation is controversial.⁶⁰

Second, on an alternative (and, to our minds, more plausible) picture of what it is to internalize a norm, it is not so far-fetched to think that complying with a norm would require following it. For example, a more cognitivist picture might see someone's internalization of a norm as equivalent to treating the norm "as reason-giving in one's practical reasoning and planning concerning some relevant circumstances."⁶¹ On such a cognitivist construal, internalization is both less mysterious and more routine.

Third, and most directly, BEGS's conclusions about the following/conforming distinction ring false. If BEGS's arguments are sound, then there should not be situations where complying with a norm requires following it. Yet complying with certain norms seems to require following them. One breaches these norms by behaving and deliberating in the ways that these norms require, but doing so for inappropriate reasons. Some argue that following is the only way to comply with certain moral norms, such as norms regarding the reparation of wrongdoing.⁶² In the legal domain, norms regarding public offices seem capable of satisfaction only through following, as we discuss more extensively in Part III.⁶³

In sum, *Explaining Norms* offers a promising definition of what norms are. The "norms as normative attitudes" view sees norms as clusters of normative attitudes and non-normative beliefs among a population. On this view, the core point of norms is to foster accountability among the members of a group. The norms as normative attitudes view does not explain norms entirely in terms of non-normative phenomena, and it generalizes to describe legal, moral, and social norms (although it also identifies important differences

59. *Id.* at 198.

60. *Id.* at 22, 68; see also Mary Clayton Coleman, *Directions of Fit and the Humean Theory of Motivation*, 86 AUSTRALASIAN J. PHIL. 127 (2008); Joshua May, *Because I Believe It's the Right Thing to Do*, 16 ETHICAL THEORY & MORAL PRACT. 791 (2013); R. JAY WALLACE, *NORMATIVITY AND THE WILL: SELECTED ESSAYS ON MORAL PSYCHOLOGY AND PRACTICAL REASON* 15-42 (2006).

61. MICHAEL BRATMAN, *FACES OF INTENTION* 197 (1999); see also Ruth Chang, *Commitments, Reasons, and the Will*, in 8 OXFORD STUDIES IN METAETHICS 74 (2013); Samuel Scheffler, *Valuing*, in REASONS AND RECOGNITION: ESSAYS ON THE PHILOSOPHY OF T.M. SCANLON (Samuel Freeman, Rahul Kumar, & R. Jay Wallace eds., 2011).

62. See Margaret Urban Walker, *The Expressive Burden of Reparations: Putting Meaning into Money, Words, and Things*, in JUSTICE, RESPONSIBILITY AND RECONCILIATION IN THE WAKE OF CONFLICT 205-25 (Alice MacLachlan & Allen Speight eds., 2013). More broadly, Richard Holton contends that the Knobe effect in experimental philosophy can be explained by the notion that intentionally conforming to a moral norm requires "be[ing] counterfactually guided by it." Richard Holton, *Norms and the Knobe Effect*, 70 ANALYSIS 417, 418 (2010). This counterfactual guidance resembles following a norm, at least on the more cognitivist understanding of that term described above.

63. See *infra* notes 116-30 and accompanying text.

among these kinds of norms). As we show in Part II, each of these features contrasts with “reductive behaviorism,” the predominant definition of norms among legal scholars.

Explaining Norms also offers insight into two aspects of norm compliance. Some norms are deliberation-sensitive. Complying with these norms turns on what goes on inside someone’s head, and not merely how she behaves. Further, there is a difference between following a norm and merely conforming to it. Although we disagree with BEGS on which types of norms are deliberation-sensitive and how to draw the following/conforming distinction, both of these notions seem to be important aspects of the normative world. Yet, as we discuss in Part III, reductive behaviorism deems both of these notions to be either unintelligible or unimportant.

II. REDUCTIVE BEHAVIORISM

Among legal scholars, the predominant definition of norms is a position that we call “reductive behaviorism.”⁶⁴ We first provide a synopsis of reductive behaviorism, then explain some of the methodological advantages of this definition, and finally examine the influence of this definition on two prominent legal scholars of norms.

Reductive behaviorism starts from the rational choice framework that BEGS adopt. However, reductive behaviorism conflicts with each of the aspects of BEGS’s norms as normative attitudes view described above. First, reductive behaviorism is reductive because it explains the acceptance of a norm in terms of non-normative considerations.⁶⁵ Second, it is behaviorist because it sees regularities of behavior as the bedrock of what norms are.⁶⁶ In BEGS’s terms, patterns of behavior determine whether a normative principle has been accepted within a group. To be sure, not all who adopt reductive behaviorism see behavioral regularities as sufficient to establish a norm. Many theorists impose additional requirements, such as that the regularity be enforced through sanctions,⁶⁷ that

64. To be sure, not all legal scholars (or even legal economists) adopt a reductive behaviorist definition of norms. Robert Cooter’s extensive writings on norms are a prominent counterexample. See Robert D. Cooter, *Three Effects of Social Norms on Law: Expression, Deterrence, and Internalization*, 79 OR. L. REV. 1 (2000); Robert D. Cooter, *Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant*, 144 U. PA. L. REV. 1643 (1996). See also Melvin A. Eisenberg, *Corporate Law and Social Norms*, 99 COLUM. L. REV. 1253 (1999); Brian Sheppard, *Norm Supercompliance and the Status of Soft Law*, 62 BUFF. L. REV. 787 (2014); Rebecca Stone, *Economic Analysis of Contract Law from the Internal Point of View* (Aug. 23, 2012) (unpublished manuscript, on file with authors).

65. NORMS, *supra* note 2, at 4.

66. See, e.g., Richard H. McAdams & Eric B. Rasmusen, *Norms and the Law*, in HANDBOOK OF LAW AND ECONOMICS 1573, 1576 (A. Mitchell Polinsky & Steven Shavell eds., 2007) (“All contributors to the [law and economics] literature seem to agree that a norm at least includes the element of a behavioral regularity in a group”) (hereinafter McAdams & Rasmusen, *Norms and the Law*); Eric Talley, *Disclosure Norms*, 149 U. PA. L. REV. 1955, 1962-63 (2001) (norms are “commonly shared practices that people tend to pursue even though legal rules alone do not compel them to do so”); ELLICKSON, *supra* note 1, at 126 (the term “[n]orm denotes both behavior that is normal, and behavior that people should mimic to avoid being punished.”) Not all views that reduce norms to non-normative phenomena construe behaviors to be bedrock. In other words, some reductionists are not behaviorists. See, e.g., Cristina Bicchieri & Ryan Muldoon, *Social Norms*, in THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2011), available at <http://plato.stanford.edu/archives/spr2011/entries/social-norms/> (“A norm cannot simply be identified with a recurrent, collective behavioral pattern.”).

67. See, e.g., Dotan Oliar & Christopher Sprigman, *There’s No Free Laugh (Anymore): The Emergence of Intellectual Property Norms and the Transformation of Stand-Up Comedy*, 94 VA. L. REV. 1787, 1812 (2008) (“To be a norm rather than a mere behavioral regularity, the rule . . . must be enforced; that is, violations must be punished.”); Claire A. Hill, *The Law and Economics of Identity*, 32 QUEEN’S L.J. 389, 414 (2007) (norms are “behavioral regularities enforced either externally, by a community, or internally, or both”); Ernst Fehr & Simon

deviations from the regularity generate disapproval,⁶⁸ or that the regularity be internalized by a wide swath of a population.⁶⁹ However, reductive behaviorism implies that, because norms are (at least) clusters of behavior, complying with a norm is a matter of behaving in a particular way.⁷⁰ Third, the core function of norms is to resolve coordination or cooperation problems.⁷¹ Fourth, reductive behaviorism explains social norms only, and is not claimed to explain either moral or legal norms.⁷²

Reductive behaviorism has several analytic advantages. It allows researchers to create testable hypotheses concerning how people behave, and so is (in theory) capable of describing how norms operate in a variety of real-world contexts.⁷³ Where manipulations of the salience of a norm are associated with changes in behavior, reductive behaviorism supports the inference that the former caused the latter.⁷⁴ Moreover, if norms are regularities of behavior, then they can be identified inductively through examining the patterns of behavior in a particular society.⁷⁵ This capacity to infer the content of norms is particularly advantageous in settings where a researcher cannot survey members of a normative community directly or where members are unlikely to have conscious access to the content of a norm.

Although many legal scholars accept reductive behaviorism uncritically or as part of disciplinary assumptions from economics, some offer a more sustained argument for

Gächter, *Fairness and Retaliation: The Economics of Reciprocity*, 14 J. ECON. PERSP. 159, 166 (2000) (a social norm is “a behavioral regularity[] that is . . . based on a socially shared belief of how one ought to behave [and] which triggers . . . the enforcement of the prescribed behavior by informal social sanctions.”); Alex C. Geisinger, *A Group Identity Theory of Social Norms and Its Implications*, 78 TUL. L. REV. 605, 608 (2004) (defining a norm as “a behavioral rule supported by a pattern of informal sanctions”); Lior Jacob Strahilevitz, *How Changes in Property Regimes Influence Social Norms: Commodifying California’s Carpool Lanes*, 75 IND. L.J. 1231, 1234 n.11 (2000) (norms are “patterns of behavior that are widely adhered to by some group of individuals, at least in part because of social pressures to conform to that norm. . . . Norms can be distinguished from other forms of human behavior that are widely adhered to for reasons having nothing to do with social pressures.”).

68. See Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338, 357 (1997) (hereinafter McAdams, *Origin*).

69. See, e.g., Bryan Druzin, *Law, Selfishness, and Signals: An Expansion of Posner’s Signaling Theory of Social Norms*, 24 CAN. J. L. & JURIS. 5, 13 (2011) (“[A] norm is . . . a pattern of behavior that has been commonly internalized to some degree or another, generating a sense of implicit obligation.”); Robert B. Ahdieh, *The Visible Hand: Coordination Functions of the Regulatory State*, 95 MINN. L. REV. 578, 641 (2010) (“Social norms, of course, are grounded in the collective practice of some regularity of behavior. Such a regularity becomes a norm, in turn, where it is followed with some sense of obligation.”).

70. See, e.g., ERIC A. POSNER, *LAW AND SOCIAL NORMS* 24 (2000) (“[S]ocial norms are always about observed behavior.”).

71. See, e.g., Paul G. Mahoney & Chris William Sanchirico, *Norms, Repeated Games, and the Role of Law*, 91 CAL. L. REV. 1281 (2003).

72. See, e.g., McAdams & Rasmusen, *Norms and the Law*, *supra* note 66, at 1577; POSNER, *supra* note 70, at 25.

73. See, e.g., POSNER, *supra* note 70, at 36-38; Robert E. Scott, *The Limits of Behavioral Theories of Law and Social Norms*, 86 VA. L. REV. 1603 (2000).

74. See, e.g., Cass R. Sunstein, *Empirically Informed Regulation*, 78 U. CHI. L. REV. 1349, 1408 (2011) (citing P. Wesley Schultz et al., *The Constructive, Destructive, and Reconstructive Power of Social Norms*, 18 PSYCH. SCI. 429, 432-33 (2007)).

75. Tanina Rostain, *Educating Homo Economicus: Cautionary Notes on the New Behavioral Law and Economics Movement*, 34 LAW & SOC’Y REV. 973, 993 (2000) (describing Eric Posner’s view as “a shift away from a deductive model of social norms in the direction of empirically grounded explanations”); Michael P. Vandenbergh, *Order Without Social Norms: How Personal Norm Activation Can Protect the Environment*, 99 NW. U. L. REV. 1101, 1114 (2005) (advocating an “inductive approach that tailors norms theory to specific clusters of behaviors and periods of time, and allows for rigorous testing of its predictive capacity for these behaviors and time periods”).

this definition. One example is Eric Posner, whose “signaling theory” aims to describe the social function that norms serve. According to Posner, social norms provide an avenue for someone to signal to others what type of person he is. Acting in accordance with a norm indicates that he is likely to be cooperative, while non-conforming behavior indicates that he is likely to be an opportunist.⁷⁶ In order to facilitate this signaling, norms must impose costly requirements on those who follow them or those who punish defections (or both) and apply to “appearance and observable behaviors” that are arbitrary (that is, ones that do not come pre-equipped with a salience that would indicate prospective cooperativeness or opportunism).⁷⁷ Posner defines social norms as “the behavioral regularities that occur in equilibrium when people use signals to show that they belong to the good type.”⁷⁸ Posner sees a necessarily tight connection between the signaling function of norms and the reductive definition of norms as behavioral regularities—if norms were not patterns of observable behaviors, then people could not use them to signal their status as cooperators through obedience or defiance.

Richard McAdams’s “esteem theory” also rests on a reductive, behaviorist definition of norms. Like Posner, McAdams’s primary ambition is to explain how norms originate and why people behave in accordance with them. McAdams starts from the assumption that everyone has a preference for esteem—that is, a desire for others to think well or not think badly of her.⁷⁹ If there is a consensus about whether “some behavior X” is esteem-worthy or not, and there is some inherent risk that others will detect whether someone X’s, and the existence of this consensus and this risk are well-known, then a norm regarding X-ing can arise if both “the consensus is that X deserves esteem” and “if the esteem benefits exceed, for most people, the costs of engaging in X.” Likewise, “if the consensus condemns X, a norm will arise if, for most people, the esteem costs exceed the benefits of engaging in X.”⁸⁰ Esteem-based sanctions can, in turn, sustain a norm if there is a widespread practice of esteeming those who do X and disesteeming those who do not do X.⁸¹

McAdams sees the esteem theory as consistent with the notion that norms are “behavioral regularities supported at least in part by normative attitudes.”⁸² On this definition, a norm is a pattern of behavior that is generally exhibited by individuals in a population; for which conformity is approved and/or nonconformity is disapproved; and for which these patterns of approval and disapproval “help[] to ensure” (in the sense of being causally implicated in) the existence of the regularity.⁸³ In other words, the normative attitudes serve as “normative incentives,” or internal and external sanctions that influence behavior

76. POSNER, *supra* note 70, at 18.

77. *Id.* at 22.

78. *Id.* at 34.

79. McAdams, *Origin*, *supra* note 68, at 356-58.

80. *Id.* at 358.

81. *Id.* at 372-73 (describing a “secondary consensus” concerning condoning of norm violations that arises if “there is some risk that one’s condoning of norm violators will be discovered without others bearing monitoring costs,” and “this risk and the secondary consensus are sufficiently well known”).

82. McAdams & Rasmusen, *Norms and the Law*, *supra* note 66, at 1576.

83. Richard H. McAdams, *Conventions and Norms: Philosophical Aspects in INT’L ENCYCLOPEDIA OF THE SOCIAL & BEHAVIORAL SCIENCES* 2735, 2739 (Neil J. Smelser & Paul B. Baltes eds., 2001). These latter two elements distinguish norms from conventions. *Id.* at 2739; McAdams & Rasmusen, *Norms and the Law*, *supra* note 66, at 1576. According to McAdams, a behavioral regularity for which normative attitudes are not present or play no causal role in its persistence is a convention, rather than a norm.

in relation to a norm.⁸⁴ On the esteem theory, the relevant normative attitudes or incentives concern esteem—that is, the desire to have others think well and not think badly of oneself, which is an instance of the broader desire for “respect or prestige.”⁸⁵

At first glance, the esteem theory does not appear to utilize a reductive definition of norms. McAdams defines norms utilizing both a non-normative element (i.e., behavioral regularities) and a normative element (i.e., normative attitudes). This definition of norms would not seem to be reductive (in the sense that BEGS use that term) because it does not characterize the facticity of norms in “purely non-normative terms.”⁸⁶ However, despite McAdams’s description of his view as non-reductive,⁸⁷ we think that the esteem theory utilizes a reductive definition of norms for the reasons discussed in the following footnote.⁸⁸

Regardless of whether it is purely reductive, the esteem theory clearly sees behavioral regularities as the bedrock of a norm. As McAdams notes, the normative attitudes related to esteem matter not in their own right, but only insofar as they sustain behavioral regularities.⁸⁹ Moreover, the esteem theory supposes that complying with a norm is entirely a matter of observable behavior, since only such behaviors would invite the “inherent risk of detection” that McAdams sees as a precondition for a norm’s stability over time.⁹⁰

While Posner’s signaling theory and McAdams’s esteem theory differ on several key issues, both views presuppose a reductively behaviorist definition of norms. Posner clearly conceives of norms in entirely non-normative terms, and McAdams arguably does

84. McAdams & Rasmussen, *Norms and the Law*, *supra* note 66, at 1578.

85. McAdams, *Origin*, *supra* note 68, at 342.

86. NORMS, *supra* note 2, at 15.

87. See Richard H. McAdams, *Signaling Discount Rates: Law, Norms, and Economic Methodology*, 110 YALE L. J. 625, 681-87 (2001).

88. There are at least two reasons to think the esteem theory utilizes a reductive (in BEGS’s sense) definition of norms. First, attitudes regarding esteem (and the broader attitudes regarding “respect” and “prestige” of which they are a subset) are not necessarily normative attitudes. For example, Stephen Darwall argues that esteem does not have the second-personal structure characteristic of normative attitudes—it can be warranted by its object, but it cannot be demanded. STEPHEN L. DARWALL, *THE SECOND-PERSON STANDPOINT* 120-22 (2006); see also R. JAY WALLACE, *RESPONSIBILITY AND THE MORAL SENTIMENTS* 131-32 (1994) (noting that states of emotions like esteem do not essentially implicate moral concepts because they “are not the sorts of states that can directly be controlled by the reasons expressed in moral principles”). These classificatory arguments have intuitive force. Many judgments of esteem do not implicate normative appraisals whatsoever—it seems routine to think well of someone (for example, because she has clear skin) without judging that she acted in the way that she ought to have acted. DARWALL, *supra*, at 122. Further, the judgment that someone acted (or did not act) in the way that she ought to have acted routinely has no effect on our relative esteem of her. If esteem is not an essentially normative attitude, then the esteem theory utilizes a reductive definition of norms because it defines norms in terms of two non-normative phenomena—behaviors and non-normative attitudes related to esteem.

A second reason to think that the esteem theory utilizes a reductive definition of norms is that McAdams does not attribute an essential role to the esteem-based attitudes in supporting behavioral regularities. Rather, it is the (non-normative) desires to obtain the esteem of others that sustain the regularities. See McAdams, *Origin*, *supra* note 68, at 342; McAdams & Rasmussen, *Norms and the Law*, *supra* note 66, at 1588. These desires might reference normative attitudes of others, but they are not essentially normative concepts. (Recall that, on BEGS’s typology, accounts that construe norms in terms of desires are reductionist. NORMS, *supra* note 2, at 22.) Thus, the best characterization of McAdams’s position is that norms are behavioral regularities sustained by an essentially non-normative desire for respect or prestige. If this interpretation is correct, then the esteem theory utilizes a reductive definition of norms because it explains the facticity of norms in terms of two non-normative phenomena—namely, patterns of behavior and non-normative desires.

89. McAdams, *Origin*, *supra* note 68, at 381.

90. *Id.* at 361; see also Philip Pettit, *Virtus Normativa: Rational Choice Perspectives*, 100 ETHICS 725, 743-44 (1990) (describing publicity assumption of norms).

too. Both theories see behaviors as the bedrock of norms. Both see the core function of norms as resolving coordination and cooperation problems, although each theory conceives of these problems somewhat differently.⁹¹ Finally, both theories aim to describe social norms, rather than norms generally.⁹² Moreover, aside from some prominent exceptions,⁹³ most legal scholars of norms accept each of the main tenets of reductive behaviorism.

III. WHAT REDUCTIVE BEHAVIORISM MISSES

Reductive behaviorism is an inadequate definition of norms. More importantly, it does not and cannot meet the goal of describing how norms work. Reductive behaviorism is committed to denying that norms could ever be deliberation-sensitive and that there could be any meaningful difference between following and conforming to a norm. Yet, as BEGS contend in *Explaining Norms* and we elaborate in this part, some norms *are* deliberation-sensitive and there *is* sometimes a meaningful difference between conforming to a norm and following it. The reductive behaviorist is doomed to miss these important aspects of the normative world.

A. *Reductive behaviorism cannot explain deliberation-sensitive norms*

According to reductive behaviorism, complying with norms is entirely a matter of behavior.⁹⁴ How someone deliberates in relation to a norm does not matter directly, although it might matter indirectly insofar as someone's deliberation predicts her behavior. Thus, reductive behaviorism is logically committed to denying the deliberation-sensitivity of norms. If compliance with a norm turned on how one deliberated, then reductive behaviorism would be incorrect because behavior would not be the bedrock of a norm.

As noted above, BEGS contend that some norms are deliberation-sensitive in that they impose requirements on how deliberation should proceed.⁹⁵ If some norms are deliberation-sensitive, then reductive behaviorism reaches incorrect conclusions about what counts as complying with these norms. To illustrate this difficulty, consider the following example.

Lucky Dentist: Dylan is a dentist who is scheduled to extract the right maxillary first premolar from his patient, Pete. To break up the monotony of his surgical life, Dylan will count the number of syllables in the first sentence that his assistant says to him and extract the tooth that corresponds (under the universal system of notation) with that number. Dylan's assistant says "How's it going, boss?", which has five syllables. Dylan commits to extracting the tooth that corresponds to the number five, which is the right maxillary first premolar. Dylan then extracts

91. POSNER, *supra* note 70, at ch. 1; McAdams & Rasmussen, *Norms and the Law*, *supra* note 66, at 1586-88.

92. POSNER, *supra* note 70, at 46; McAdams & Rasmussen, *Norms and the Law*, *supra* note 66, at 1577.

93. *See supra* note 64.

94. *See supra* note 66.

95. *See supra* note 31 and accompanying text.

Pete's right maxillary first premolar.

Does Dylan live up to the norms for dentists? We think not. Yet this verdict can't be explained by Dylan's observable behavior: Dylan extracts Pete's right maxillary first premolar, which is what the norms for dentists would seem to require.⁹⁶ Rather, the problem is with Dylan's deliberation. Complying with the norms for dentists requires the interests of the patient to figure into the dentist's practical deliberation in certain ways. Yet Pete's interests do not figure into Dylan's decision about what to do. In determining which of Pete's teeth to extract, what matters to Dylan is the number of syllables in his assistant's sentence, not which tooth's extraction would most help Pete. It is an accident that Dylan's behavior winds up serving Pete's best interests.

Reductive behaviorism gets the wrong answer in the Lucky Dentist case. The reductive behaviorist seemingly must embrace the conclusion that Dylan complies with the norms for dentists because, in extracting Pete's right maxillary first premolar, he behaves in way that these norms require. The problem is not merely that some versions of reductive behaviorism seem to reach the wrong verdict in this case. Rather, *every* version of reductive behaviorism *necessarily* reaches the incorrect verdict because of the commitment to define norms as behavioral regularities. If behavioral regularities are the bedrock of norms, then the normative principles contained in norms can only govern behaviors.⁹⁷

The Lucky Dentist case suggests that some non-formal norms (like the norms for dentists) govern both behavior and deliberation. However, contrary to BEGS's claim that formal norms are deliberation-insensitive,⁹⁸ some legal norms also impose requirements on deliberation.

Fiduciary duties are a clear example of such deliberation-sensitive formal norms. Fiduciary duties arise out of fiduciary relationships like those between attorneys and clients, corporate board members and shareholders, or trustees and beneficiaries.⁹⁹ Fiduciary duties are typically taken to include at least a duty of care and a duty of loyalty.¹⁰⁰ Both of these component duties can be formulated to concern behavior: the duty of care is often taken to be satisfied when the fiduciary behaves in the way that a prudent or diligent agent would,¹⁰¹ and some see the duty of loyalty as satisfied when the fiduciary behaves in accordance with the "no conflict" rule (which prohibits "acting with a conflict . . . [of] inter-

96. For example, Dylan's behavior does not obviously run afoul of the American Dental Association's Principles of Ethics and Code of Professional Conduct. Dylan does not violate the principle of nonmaleficence because he does not harm Pete. ADA Code, § 2. Dylan's actions also seem consistent with the principle of beneficence because his behavior has the effect of promoting Pete's welfare. *Id.* §3.

97. By contrast, an account that did not see behavioral regularities as the bedrock of norms could make sense of the idea that a norm would regulate something other than behavior. For example, on the norms as normative attitudes view that BEGS advance, a norm could regulate deliberation if the normative principle contained in the norm calls for deliberating in a certain way, and the principle-corresponding normative attitudes would attach if someone did not deliberate in the way that the principle required.

98. NORMS, *supra* note 2, at 50-51.

99. Robert Flannigan, *The Fiduciary Obligation*, in 9 OXFORD J. LEGAL STUD. 285, 286 (1989).

100. Arthur B. Laby, *The Fiduciary Obligation as the Adoption of Ends*, 56 BUFF. L. REV. 99, 105-06 (2008).

101. TAMAR FRANKEL, FIDUCIARY LAW 169 (2011) ("The duty of care requires fiduciaries to execute their services, and execute them well.").

est”) and the “no profit” rule (which prohibits “making a profit [off] [] the fiduciary position”).¹⁰²

In addition to these behavioral requirements, fiduciary duties also seem to impose deliberative requirements. In some contexts, the duty of care is explicitly formulated in a way that imposes requirements on how a fiduciary deliberates. On such a construal, the “duty of care” amounts to a duty to be careful. For example, the norms governing the professional conduct of attorneys require not just that an attorney’s course of action “advance a client’s lawful objectives,” but also that it be “reasonably calculated” to do so.¹⁰³ By implication, a course of action that happened to advance the client’s lawful objectives but was not reasonably calculated to do so would run afoul of this standard. Furthermore, many scholars construe the duty of loyalty to impose deliberative requirements on fiduciaries.¹⁰⁴ According to these scholars, a fiduciary does not satisfy her duty of loyalty unless the interests or ends of the principal play a specific role in her deliberation about what to do.

Of course, someone attracted to reductive behaviorism about norms might also be attracted to the view (popular among law and economics scholars) that fiduciary duties are fully explicable in terms of behavior.¹⁰⁵ However, behaviorism about fiduciary duties generates absurd conclusions that parallel those in the Lucky Dentist case. For example, behaviorism about fiduciary duties would imply that a corporate director could satisfy her duty of loyalty by choosing a course of action on a whim (that is, without any consideration of the merits of the decision), so long as this course of action is the one that she would have chosen if she had deliberated properly.¹⁰⁶

We do not mean to suggest that all (or even most) legal or social norms are deliberation-sensitive,¹⁰⁷ nor to deny that most deliberation-sensitive norms also impose behavioral requirements. Indeed, there is good reason to think that few legal norms will be deliberation-sensitive. As BEGS note, legal norms are policed through institutions in a way that social and moral norms are not.¹⁰⁸ It is difficult to police deliberation-sensitive norms because it is difficult to determine how someone has deliberated. By contrast, it is much easier to establish when someone has behaved in a particular way. These epistemic challenges to policing deliberation might explain why deliberation-sensitive legal norms are rare, but they do not support the more ambitious claim (entailed by reductive behaviorism) that deliberation-sensitive legal norms cannot exist.¹⁰⁹

102. Matthew Conaglen, *The Nature and Function of Fiduciary Loyalty*, 121 LAW Q. REV. 452, 459-60 (2005).

103. AMERICAN LAW INSTITUTE, RESTATEMENT (THIRD) OF LAW GOVERNING LAWYERS, §16(1) (2000).

104. See, e.g., Lionel Smith, *The Motive, Not the Deed*, in RATIONALIZING PROPERTY, EQUITY, AND TRUSTS: ESSAYS IN HONOUR OF EDWARD BURN (Joshua Getzler ed., 2003); Laby, *supra* note 100; Galoob & Leib, *supra* note 43.

105. See generally, Frank H. Easterbrook & Daniel R. Fischel, *Contract and Fiduciary Duty*, 36 J. L. & ECON. 425 (1993); Larry E. Ribstein, *Fencing Fiduciary Duties*, 91 BOSTON U. L. REV. 899 (2011).

106. Cf. *In re Walt Disney Co. Derivative Litig.*, 825 A.2d 275, 289 (2003) (stating that a fiduciary duty would be violated if corporate directors “consciously and intentionally disregarded their responsibilities” by “adopting a ‘we don’t care about the risks’ attitude concerning a material corporate decision”).

107. On the other hand, there is considerable disagreement among philosophers about whether moral norms are necessarily (or ever) deliberation-sensitive. See S. Matthew Liao, *Intentions and Moral Permissibility: The Case of Acting Permissibly with Bad Intentions*, 31 LAW & PHIL. 703 (2012).

108. NORMS, *supra* note 2, at 43-46.

109. The deliberation-sensitivity of norms is worthy of future study. It seems possible that groups could cus-

In sum, some social norms are deliberation-sensitive. Complying with these norms requires both behaving and deliberating in specific ways. Moreover, the example of fiduciary duties indicates that, *pace* BEGS, some legal norms are deliberation-sensitive. One can violate a fiduciary duty when she deliberates in a particular way (or fails to), even if she otherwise behaves in the way that a careful and loyal person would behave. Reductive behaviorism cannot explain the existence of deliberation-sensitive norms; indeed, it sees the possibility of such norms as paradoxical. Thus, reductive behaviorism obscures part of the normative world.

B. Reductive behaviorism cannot explain the following/conforming distinction.

Reductive behaviorism also cannot make sense of the difference between conforming to and following a norm. Recall that, on BEGS's definition, one follows a norm by acting in the way that the norm requires and because the norm requires it. By contrast, one conforms to a norm when she acts in the way that the norm requires while taking the norm to be an instrumental reason for action.¹¹⁰ The reductive behaviorist would be skeptical that following could ever be the requisite mode of responding to a norm. To be sure, many legal scholars (including some committed to reductive behaviorism) devote significant attention to the question of how people do and might internalize norms.¹¹¹ On reductive behaviorism, however, internalization only matters insofar as it generates or supports regularities in behavior. Norms that are internalized by a population are more likely to generate and sustain these regularities than norms that are not internalized. Thus, at the macro level, internalization might be useful to ensure the stability of norms, but it does not matter in its own right. One need not internalize any norm in order to comply with it.¹¹²

This skepticism is unwarranted. Complying with some norms requires following them. Our contention here runs contrary to both reductive behaviorism and the conclusions of BEGS, who deny that any moral or legal norms must (or even could) be widely followed and see conformity with social norms to be as good as following them.¹¹³ In short, BEGS draw the following/complying distinction in the wrong place, and the reductive behaviorist cannot appreciate the importance of this distinction at all.

Under what conditions would a norm call for following as the paradigmatic or requisite mode of compliance? Assume that one can internalize a norm by taking it to provide

tomize the deliberation-sensitivity norms. To put this point in BEGS's terms, a norm embodying the same normative principle might be deliberation-sensitive in one group and deliberation-insensitive in another group. Likewise, the degree to which a norm is deliberation-sensitive might vary among groups. We conjecture that deliberation-sensitivity reflects a broader concern with what might be called "non-accidentalness," which is in turn a way to elevate or reflect the importance of a normative principle within a group.

110. NORMS, *supra* note 2, at 195.

111. See, e.g., Cooter, *supra* note 64; Yuval Feldman & Robert J. MacCoun, *Some Well-Aged Wines for the "New Norms" Bottles: Implications of Social Psychology for Law and Economics*, in THE LAW AND ECONOMICS OF IRRATIONAL BEHAVIOR (Francisco Parisi & Vernon L. Smith eds., 2005); Robert C. Ellickson, *The Evolution of Social Norms: A Perspective From the Legal Academy*, in SOCIAL NORMS 35 (Michael Hechter & Karl-Dieter Opp eds., 2001); Richard H. McAdams, *Resentment, Excuse, and Norms*, in THE HART-FULLER DEBATE IN THE TWENTY-FIRST CENTURY 249, 253-54 (Peter Cane ed., 2010).

112. For example, McAdams contends that both a sense of obligation and concern for esteem provide adequate incentives to comply with a norm. *Id.* at 256-57; Richard H. McAdams, *Conformity to Inegalitarian Conventions and Norms: The Contribution of Coordination and Esteem*, 88 THE MONIST 238, 251-52 (2005).

113. See NORMS, *supra* note 2, at 195.

a non-instrumental reason to act in accordance with the norm.¹¹⁴ This notion of internalization requires a certain conscientiousness regarding the norm. Someone who justifies her actions based on first-order considerations without taking the norm itself to have practical significance would fail to internalize (and, by extension, would not follow) the norm.

It strikes us that complying with many social norms requires this kind of conscientiousness. If so, then following these norms is the paradigmatic mode of complying with them. For example, take the norm of spousal fidelity.¹¹⁵ This norm has an obvious behavioral component: to be faithful, one must, *inter alia*, avoid engaging in sexual behavior with people who are not one's spouse. Suppose that Ann behaves in the way that the norm of spousal fidelity requires, but sees this norm as providing merely instrumental reasons for action. Ann is disposed to engage in dalliances with those other than her spouse, but believes that she should avoid this behavior because, if detected, it would invite unwelcome disapprobation from others. In many communities, it might be judged that Ann fails to comply with the norm of spousal fidelity. If so, then in these communities complying with the norm of spousal fidelity would require following it.

Some legal norms, in particular norms related to public offices, also call for following. Many public offices require their inhabitants to have or maintain higher-order commitments regarding the offices themselves and the systems within which the offices operate. Someone can fail to comply with the norms of these public offices when she lacks these commitments, even when her behavior and deliberation otherwise mirror that of a committed office-holder.

To illustrate this phenomenon, consider an example based on the television series *The Americans*.¹¹⁶ As background, *The Americans* is set during the waning days of the Cold War in the early 1980s. Philip and Elizabeth Jennings appear to be an attractive married couple living in a suburb of Washington, D.C. However, Philip and Elizabeth are elite Soviet spies who have operated under deep cover in the United States since the late 1960s. The external trappings of Philip's and Elizabeth's lives are an artifice that allows them to pursue their true mission. Against this backdrop, consider the following scenario.

Undercover Judge: As part of a KGB plot, Elizabeth undertakes a legal education and has a successful legal career. After several years of distinguished practice, she is appointed as a federal judge. The KGB approves this appointment in order to allow Elizabeth to use her judicial discretion to subtly subvert the American legal system from within whenever possible. In order to instruct Elizabeth on how to decide the cases before her and to evade detection, the KGB (unbeknownst to Elizabeth) kidnaps another judge. The kidnapped judge is forced to review the pleadings and transcripts of oral arguments in all cases before Elizabeth, as well as to write the legal opinions that he would write if he

114. As would be consistent with the cognitivist picture of internalization described *supra*, at note 61. As noted above, this standard of internalization is less demanding than the Humean standard that BEGS posit, on which internalizing a norm requires forming non-derivative desires regarding the norm. NORMS, *supra* note 2, at 201-02.

115. David Shoemaker, *On Criminal and Moral Responsibility*, in 3 OXFORD STUDIES IN NORMATIVE ETHICS 154, 173 (Mark Timmons ed., 2014).

116. *The Americans* (FX Television Broadcast, 2013).

were the judge in these matters. These opinions are then transmitted to Elizabeth through her KGB handlers. Elizabeth then promulgates these opinions as if she had written them.

Assume that the norms for judges are codified by the American Bar Association's Model Code of Judicial Conduct.¹¹⁷ The Model Code consists of both canons, or "overarching principles of judicial ethics that all judges must observe,"¹¹⁸ and specific rules (derived from each canon) that govern the actions of judges. Canon 1 states that a judge must "uphold and promote the independence, integrity, and impartiality of the judiciary, and . . . avoid impropriety and the appearance of impropriety."¹¹⁹ Canon 2 states that a judge must "perform the duties of judicial office impartially, competently, and diligently."¹²⁰ Some of the rules based on Canon 2 impose deliberative requirements. For example, Rule 2.2 says that a judge must "uphold and apply the law" and "perform all duties of judicial office fairly and impartially."¹²¹ Impartiality, in turn, requires the judge to be "objective and open-minded."¹²² Other rules impose behavioral requirements. For example, Rule 2.5 establishes that competence and diligence "requires the legal knowledge, skill, thoroughness, and preparation reasonably necessary to perform a judge's responsibilities of judicial office."¹²³

In *Undercover Judge*, does Elizabeth comply with the norms for judges?¹²⁴ We think not. Elizabeth conforms to these norms because her behavior and deliberation meet the standard of Canon 2. Every ruling Elizabeth makes is one that a competent judge would make. Every opinion she promulgates is both based on the relevant legal materials and identical to what a competent judge would write. However, Elizabeth does not follow the norms for judges. Regardless of whether Elizabeth's actions have the tendency to promote the "independence, integrity, and impartiality of the judiciary," she does not *uphold* these values, as Canon 1 requires. Another way to state the problem is that Elizabeth lacks a higher-order commitment to the legal system that judges must have in order to comply with the norms for judges. Elizabeth treats the norms for judges as providing purely instrumental reasons for her to act: her pattern of behavior and deliberation is consistent with the norms, but she behaves and deliberates in this way in order to evade detection and

117. MODEL CODE OF JUDICIAL CONDUCT (2007). The Model Code would meet BEGS's definition of a formal norm, since it is promulgated by an identifiable source (the American Bar Association) that is taken to have the authority to interpret and revise these norms.

118. *Id.* at Scope ¶ 2.

119. *Id.* at Canon 1.

120. *Id.* at Canon 2.

121. *Id.* at Rule 2.2.

122. *Id.* at Rule 2.2, cmt. 1.

123. MODEL CODE, Rule 2.5, cmt. 1.

124. Whether Elizabeth complies with the norms for judges differs from whether her opinions are law, or whether a system in which a significant portion of officials took Elizabeth's position would be a legal system at all. These latter two questions are debated by jurisprudential scholars grappling with H.L.A. Hart's contention that judges and other officials in a legal system must take the "internal point of view" regarding fundamental norms. Compare H. L. A. HART, THE CONCEPT OF LAW 91 (2d ed. 1997); Scott J. Shapiro, *What is the Internal Point of View*, 75 FORDHAM L. REV. 1157 (2006); Christopher Kutz, *The Judicial Community*, 35 NOÛS 442 (2001); with Kenneth M. Ehrenberg, *The Anarchist Official: A Problem for Legal Positivism*, 36 AUSTRALIAN J. LEGAL PHIL. 89 (2011). In saying that Elizabeth does not comply with the norms of judging in *Undercover Judge*, we take no stance on whether her decisions are law.

further the interests of her Soviet principals. Elizabeth is not committed to these broader values, and so she does not uphold them. This lack of commitment explains why Elizabeth breaches the norms for judges.

Reductive behaviorism reaches a different conclusion in the Undercover Judge case. If norms are clusters of behaviors, then complying with a norm cannot require anything more than behaving in accordance with the norm. Because Elizabeth behaves in exactly the ways that the norms for judges require, the reductive behaviorist must conclude that she complies with these norms. This conclusion is both intuitively implausible and inconsistent with the Model Code. Therefore, reductive behaviorism does not (and cannot) describe what it would take to comply with the norms for judges.

BEGS also might contend that Elizabeth has complied with the norms for judges in Undercover Judge, although the norms as normative attitudes view does not bind them to this conclusion.¹²⁵ The norms for judges are formal norms, and BEGS contend that following cannot be the paradigmatic mode of complying with formal norms. This conclusion is both intuitively implausible and in tension with the Model Code.

BEGS might respond by arguing that the Undercover Judge case proves too much. On BEGS's definition of internalization, it is arguable that any judge who acts in order to uphold some purpose broader than the norms for judging would fail to follow these norms. Yet not every judge who acts with a broader purpose would seem to breach the norms for judging in the way that Elizabeth does. For example, a judge who made his decisions in order to promote the American way of life would, on BEGS's definition, fail to internalize the norms for judges because he would treat these norms as instrumentally important (that is, important insofar as they tend to promote the American way of life). Because one must internalize a norm in order to follow it, such a judge would not follow the norms for judges. Yet, intuitively, such a judge would not seem to violate the norms for judges.

This response rests on BEGS's Humean account of internalizing a norm, which we take issue with above.¹²⁶ In addition, the notion of internalization in this imagined response seems too narrow. One could seemingly internalize a norm not only by taking the norm itself to be a reason for action, but also by taking a broader principle (into which the principles contained in the norm figure inextricably) to be a reason for action. If so, then the judge who is committed to the American way of life would count as internalizing the norm for judges to the extent that he saw these norms as integral to the American way of life.

To be sure, norm following is a tricky topic. Norms for which following is the paradigmatic or requisite mode of compliance should be rare. However, while BEGS would explain this rarity in terms of the irrationality of following a legal norm, we think that there is a more straightforward explanation in terms of political liberalism. It is difficult to determine when someone has followed a particular norm. If a significant portion of legal norms required following, then the question of whether someone had complied with a legal norm would routinely remain unresolved. This would violate a number of liberal principles of legality that govern the enforcement of legal norms, including the requirements that the

125. On the norms as normative attitudes view, a norm could require following if the normative principle contained in the norm included a self-reference, and if the principle-corresponding normative attitudes that constituted the norm attached to efforts that did not satisfy this self-referential requirement.

126. See *supra* notes 57-61 and accompanying text.

law be public and stable.¹²⁷

Moreover, we conjecture that, where a norm calls for following, a wide variety of mental states will likely suffice.¹²⁸ In the Undercover Judge case, the commitments that Elizabeth lacks seemed just as problematic as the ones that she has. Thus, a broad range of intentional stances seem compatible with someone's following the norms of an institutional role. However, some commitments are bedrock. One who lacks these bedrock commitments does not live up to the norms of that role.

It might be argued that we have based a broad conclusion about norm following on a stylized example. Here is a positive claim that strikes us as plausible, but for which space constraints preclude an adequate defense: following is likely to be the requisite mode of complying with a norm whenever it is meaningful to criticize someone for acting in bad faith with regard to that norm. This conjecture reconciles the analysis of norm following with the earlier discussion of the deliberation-sensitivity of certain norms. When it is possible to criticize someone for acting in bad faith, a norm probably has some deliberation sensitivity. Therefore, deliberation-sensitive norms are the norms that are most likely to call for following. In these cases, the requirement that someone follow the norm is a way of ensuring the appropriate connection between her deliberation and her behavior. Dylan's actions in Lucky Dentist and Elizabeth's actions in Undercover Judge are both arguably examples of bad faith, and both cases also seem to involve inappropriate patterns of deliberation. This intertwining of deliberation-sensitivity and norm following is also captured by recent attempts to apply fiduciary concepts to judges and other public officials.¹²⁹ The agent's discretionary authority and capacity for self-dealing, coupled with the difficulties of the monitoring how the agent exercises this authority, lead to worries about how the agent behaves, how she deliberates, and why she decides to behave and deliberate in the ways that she does.¹³⁰ However, reductive behaviorism rules out in advance the possibility that bad faith could matter in connection with any norm. Thus, reductive behaviorism occludes how at least some norms operate.

CONCLUSION

BEGS's insights about the nature of norms and compliance are revelatory, even though we disagree with some of the conclusions that they reach. *Explaining Norms* should prompt greater attention to foundational issues in the study of norms by legal scholars. Some norms seem to govern behavior and deliberation, and some seem to only concern behavior. It also strikes us (although not BEGS) as an open question which norms must be

127. Cf. Adam Hill, *Stability, Assurance, and the Concept of Legal Guidance*, LAW & PHIL. (forthcoming, 2015), available at <http://link.springer.com/article/10.1007%2Fs10982-014-9204-y>.

128. Similarly, Scott Shapiro reads Hart's notion of the "internal point of view" that must be taken by a critical mass of legal officials to be capacious enough to include a wide variety of mental states, so long as these officials "accept the shared master plan of the system." SCOTT SHAPIRO, LEGALITY 97, 204-05 (2011). See also MARK BEVIR, DEMOCRATIC GOVERNANCE 267 (2010) ("An interpretive social science based on recognition of local reasoning typically leads to decentering, where to decenter is to analyze a practice in terms of the disparate beliefs of the actors. . . . Decentering . . . consists in showing how an apparently monolithic institution actually embodies diverse, contested webs of belief informed by different traditions.").

129. See, e.g., Ethan Leib, David Ponet, & Michael Serota, *A Fiduciary Theory of Judging*, 101 CAL. L. REV. 699 (2012); Sung Hui Kim, *The Last Temptation of Congress: Legislator Insider Trading and the Fiduciary Norm Against Corruption*, 98 CORNELL L. REV. 845 (2013).

130. EVAN FOX-DECENT, SOVEREIGNTY'S PROMISE: THE STATE AS FIDUCIARY 107-08 (2011).

followed and which can be satisfied through conformity. But these are questions worth asking. Reductive behaviorism denies that these questions are even intelligible, let alone interesting. Thus, this position is unsound.

Why should a legal scholar (or anyone else) care about conceptual soundness? As Gerald Gaus has noted, a theory of norm-guided activity can be powerful even if it defines norms in a way that “does not perfectly track applications of the term ‘norm.’”¹³¹ This position is wise—unsoundness matters much more in the seminar room than in doing field research or making policy. However, not all conceptual flaws are easily dismissed. Reductive behaviorism is imperfect in a way that leads to an inadequate description of what it is to live under and comply with norms. Nor is this imperfection inevitable, as several legal scholars have appreciated the significance of (and attempted to model) the cognitive dimensions of norms.¹³² A theory that reduces all norms to behaviors is foreordained to miss important features of the normative world.

131. Gerald Gaus, *Review of Explaining Norms*, NOTRE DAME PHILOSOPHICAL REVIEWS (2014), available at <http://ndpr.nd.edu/news/45998-explaining-norms/>. Bicchieri also doubts the wisdom of establishing what is a “necessary and sufficient condition for being a norm,” given the “semantic vagueness surrounding the concept of norm is common to all social constructs.” Bicchieri & Muldoon, *supra* note 66.

132. *See supra* note 64.