

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Publications - Department of  
Philosophy

Philosophy, Department of

---

7-2014

**Book Review: Brennan, Geoffrey; Eriksson, Lina; Goodin, Robert E.;  
and Southwood, Nicholas. Explaining Norms.**

David K. Henderson

Follow this and additional works at: <https://digitalcommons.unl.edu/philosfacpub>



Part of the [Philosophy Commons](#)

---

This Article is brought to you for free and open access by the Philosophy, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications - Department of Philosophy by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

## BOOK REVIEWS

Brennan, Geoffrey; Eriksson, Lina; Goodin, Robert E.; and Southwood, Nicholas. *Explaining Norms*.

Oxford: Oxford University Press, 2013. Pp. 290. \$55.00 (cloth).

*Explaining Norms* is a work in philosophy of social science aspiring to provide an account of norms, their general character, their kinds (formal, legal, moral, and social), what they can explain, and what explains their dynamic (emergence, persistence, and unraveling). The authors engage with various positions in ethics, political philosophy, and (to some extent) the philosophy of law. The discussion is rewarding and inventive—it provides distinctive and intriguing views on several topics (e.g., on the distinction between moral and social norms). There are a lot of ideas here. Perhaps this is predictable, given that the work is a product of four capable minds. What is surprising is the range of ideas and arguments on which the authors manage to agree and out of which they construct one reasonably cohesive account.

Given the wide range of literatures discussed, readers are likely to find much of interest. Not surprisingly, some related literatures do seem to be underplayed—treated in a few footnotes and somewhat by the way, with little development of systematic connections. There are thriving literatures in comparative psychology/ethology, moral psychology, and cultural anthropology that are devoted to how we humans manage to cooperate and coordinate as we do. While there are footnotes to some of this literature (see the index for, e.g., Boyd, Bowles, Camerer, Gintis, and Henrich), many readers would have benefited from a discussion that more fully related the position developed in *Explaining Norms* to that work in experimental economics and anthropology. Discussing the relationships with work in moral psychology and ethology would also have been appreciated (Haidt, de Waal, and Tomasello are not mentioned). Still, there is very much to like about what is treated here.

The authors seek conceptually individually necessary and jointly sufficient conditions for being a norm. On the account provided, norms are something on the order of normative principles accepted in some group (3–4). Thus, norms involve some normative principle, possessing “a certain generality of scope and application.” These principles need not be objectively correct or fitting, and they may be objectively “simply awful” (3). The central questions have to do with what is involved in some group accepting these normative principles. The authors locate their position in two dimensions.

First, there are individualistic and holistic accounts of acceptance. Individualistic accounts hold that “to be accepted in a particular group or community

just is for certain things to be true of (some significant portion of) the *individual members* of that group” (4). In contrast, holistic accounts hold that acceptance requires that certain things be true of the group as a whole, where these cannot be reduced to things that are true of the individual members of the group. The authors develop an individualistic account.

Second, there are reductive and nonreductive accounts of acceptance. These differ over whether the idea of ‘accepting a rule or normative principle’ is itself to be understood in “non-normative or normative terms.” According to reductive accounts, for a norm to be accepted in a particular group “just is for certain non-normative truths to hold of either (some significant portion of) the individuals or the group as a whole” (4). What supposedly is it for acceptance to be a normative truth or matter? Here, the authors make much of whether the relevant folk adopt “normative attitudes,” rather than merely (for example) desiring or preferring that folk conform to the principle. Apparently, desires or preferences having to do with conformity to normative principles count as nonnormative facts or truths, while folk adopting normative attitudes in connection with that normative principle constitute normative matters or truths in the relevant sense. (For the authors, there is no suggestion here that folk who accept some normative principle in this normative sense really ought to conform to that principle. What is at stake is just whether their acceptance involves stances that are themselves contentfully richer than mere preference structures and behaviors.) While I find the authors’ way of writing of normative and nonnormative matters jarring (I would generally not think of folk adopting normative stances in connection with some principle a “normative truth”), I suppose it occasions no misunderstanding, once noted.

The authors pursue a nonreductive account of norms. They believe that their combination of individualism and nonreductionism is nonstandard. Individualism and reductive accounts are commonly allied. Holistic and nonreductive accounts are commonly allied.

As the authors understand the territory, reductive accounts tend to fall into two clusters: those that treat norms as practices (as conventions or customs) and those that treat norms as clusters of desires or preferences. A social practice is here understood as “a regularity in behavior among the members of a group that is explained, in part, by the presence within the group of pro-attitudes (or beliefs about the presence of pro-attitudes) towards the relevant behavior that are a matter of common knowledge among members of the group” (16). Paradigmatic accounts of norms as conventions understand these as equilibria behavior in a game played repeatedly—as a solution rooted in preferences and some common knowledge. Lewis advanced an early and influential version on conventions.

To show that social practice accounts are mistaken regarding the conditions sufficient for there being a norm, the authors adduce the thought experiment of Imelda’s Inn (18). Here some group of individuals, fond of each other’s company, meet regularly for noon lunch (although other times would have suited as well) at a particular place (which is not the unique convenient place). Loathing normative rules or principles, they repudiate any coordinating normative principle and would not enjoy their gathering were they to have thought others thought that their showing up was subject to any normative requirement. In-

dividuals within this lunching group would not normatively evaluate others within the group, were they not to show up. The authors insist that, with no place for normative attitudes, there would be no norm here—although there is a conventional practice.

The authors also argue that social practices are not necessary for there being a norm. In their thought experiment of the Moldovans (20), they characterize a people most of whom judge that one ought not *U* (urinate in public swimming pools), are disposed to disapprove of anyone *U*-ing, and are disposed to think such disapproval fitting. Still, most Moldovans do *U*, and most feel guilty upon *U*-ing and hide their *U*-ing. The Moldovans apparently do have a norm of not *U*-ing, although they do not have a social practice of not *U*-ing.

Other reductive accounts understand norms as “clusters of non-normative attitudes”—specifically as “clusters of desires (possibly accompanied by certain non-normative beliefs)” (22). Such norms-as-desires accounts do allow that there may be norms that are commonly violated. Thus, these accounts recognize that norms are not social practices.

The authors judge accounts in terms of conditional (rather than unconditional) desires the most promising (23–24) of these accounts. They focus on Bicchieri’s account, which they gloss this way:

**A normative principle P** is a norm within group G if a significant proportion of the members of G prefer to comply with P on the condition that

- a. A significant proportion of members of G comply with P; and
- b. Either
  - i. A significant proportion of the members of G expect her to comply with P; or
  - ii. A significant proportion of the members of G expect her to comply with P, prefer her to comply with P, and may sanction her for not complying with P. (24)

Clause (a) is probably a minor distortion, in that Christina Bicchieri (*The Grammar of Society* [New York: Cambridge University Press, 2006], 11–12) explicitly allows for norms that are not social practices. On Bicchieri’s account, clause (b) is pivotal. The central attitude invoked there is a kind of conditional preference. Desires (preferences) themselves are not normative attitudes.

The authors argue that satisfying Bicchieri’s account would not be sufficient for a group to have a norm: they consider a people, the Chastians. Among these chaste folk, significantly many actually conceive of a different, very nonchaste, principle or code, which they secretly conditionally prefer (25). As the case is developed, the nonchaste principle would constitute the Chastians’ norm on the conditional preference account of norms. Yet it seems right to say that their norms are the chaste principles, not the alternative principles.

The authors also suggest reasons to think that conditional preference accounts do not capture necessary conditions for there being norms. Among the Philanthropians, virtually all members judge that one must donate one-half of one’s income to help feed the poor, are disposed to disapprove of those donating less, and are disposed to regard such attitudes as appropriate. Yet we are asked to

suppose that these Philanthropians generally do not unconditionally desire to so behave, nor do they conditionally so prefer. They do not care what others do with their money and simply prefer to buy luxury goods themselves. As a result, they do not so donate.

The authors insist that the Philanthropians have a norm. But, one may wonder about the case itself. The Philanthropians supposedly disapprove of those that do not so donate. Thus, they think it wrong not to so donate. But they don't care what others do with their money? I have reservations about the moral and related psychology that must be supposed in this case. It is hard to make sense of folk with normative stances without having some corresponding conditional desires or preferences. It seems that the normative attitudes supposed to be had by the Philanthropians plausibly would involve some caring.

Ultimately, *Explaining Norms* advances an account of norms turning on "P-corresponding normative attitudes." These are normative attitudes "appropriately reflecting *the content and normative force of*" the relevant normative principle, P. Normative attitudes include:

- a) Normative beliefs, judgments and other cognitive states,
- b) Normative expectations,
- c) Reactive attitudes and dispositions to have (a) or (b),
- d) Any attitude entailing a–c. (29)

A normative principle P is said to be a norm within a group G if and only if:

- i. A significant proportion 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. (29)

This account of norms is worth careful reflection. It should be attractive to one who (like myself) is sympathetic to the idea that an adequate account of norms will need to turn on various normative attitudes. Arguably, much of the contemporary work in moral psychology and evolutionary game theory is devoted to understanding the character of such judgments, expectations, and reactive attitudes. Accounts of norms in terms of preferences should own up to the importance of such attitudes.

The authors advertise their account as "interesting and important" in part because it is supposed to have implications for what is the function of norms—for "what they are for" (35). Rejecting two familiar ideas, the authors argue that the function of norms is neither facilitating coordination nor facilitating cooperation. While norms do sometimes contribute to coordination and cooperation, understanding "the core function" of norms primarily in these terms is said to be a mistake (35). Their "core function" is said to be a matter of making folk accountable—where "accountability involves . . . others having a recognized right or entitlement to determine how one ought to behave" (36).

I must admit that I find the arguments here rather puzzling. It certainly can be said that, on the account advanced, norms do make for a "recognized entitlement"—as it treats "recognition" (in the sociological sense) of some applica-

tions of some normative principle by folk in the relevant group as conceptually constituting their having a norm. “Recognition” of entitlements then is constitutive of there being a norm. But, to call this “the function” of norms seems analogous to saying that the function of automobiles is automotive transportation. Further, it is somewhat obscure just how the authors would have us understand the notion of “a function.” There are two prominent accounts of functions in the philosophy of science. On the account articulated by Robert Cummins (*The Nature of Psychological Explanation* [Cambridge, MA: MIT Press, 1983]), a function is a causal relationship that features in an analysis of a (perhaps complex) disposition or capacity of a class of systems. On the account articulated by Larry Wright (*Teleological Explanation* [Berkeley: University of California Press, 1976]), a function is a causal/functional relationship featuring in a kind of selection explanation of the items involved in that relationship. However, it is plausible that coordination and cooperation facilitation (as well as accountability) constitutes both Cummins-style and Wright-style functions of norms. The fact that a significant proportion of a group “have P-corresponding normative attitudes” may feature in an analysis of a social system that accounts for a kind of equilibrium in which folk are held or judged accountable. But, depending on the relevant principle, P, the equilibrium afforded by the norm will afford a form of social coordination or cooperation that would not be incidental to the analysis or explanation.

The authors discuss various kinds of norms. Formal norms are those in which application—including punishment—is delegated to some power, office, or group. These are found fitting only when they regulate actions. Informal norms are those the application of which is not delegated, so that they must be interpreted and applied by a substantial number of the norm holders themselves.

The authors also distinguish between moral norms and social norms. Here they reject the idea that the distinction can be drawn on the basis of certain “purely formal” features of the principles: their scope, their being conditional or unconditional, or their having normativity in being “intrinsically motivating” or rooted in certain reactive attitudes (59). The authors also argue against distinguishing moral and social norms in terms of the content of the normative principles involved. Both moral and social norms can call on others to be other regarding, or impartial, or to respect individuals *qua* individuals. They insist that moral norms may lack such content features and that social norms may possess such content features. Their central positive idea is this: the motivation for the judgment is separable from the content of the judgment—and whether a judgment is a moral judgment turns on the grounds in the agents’ minds for their judgment. This suggests that we should distinguish moral norms from social norms, not by the contents of the principles involved but by the grounds featured in the minds of the norm holders (58). Moral norms are said to be “constituted by attitudes that are necessarily *practice-independent* in the sense that presumed social practices may constitute no part of their grounds” (59). Social Norms, in contrast, are said to be “constituted by attitudes that are necessarily grounded, in part, in presumed social practices” (59).

This intriguing suggestion seems on its face to have the implication that common norms for such matters as fidelity within marriage turn out to be social

norms rather than moral norms—presumably most norm holders have motivating attitudes in which the social institution and practice of marriage plays a part. Again, the authors' rich discussion is worth careful attention. They discuss cases in which a social practice may feature in the thought of the norm holders yet do so in "a wholly derivative way" (68–70). In a footnote (70 n. 36) they consider a case in which the agent takes there to be some social practice, P, takes the social practice P to make for valuable results, and takes the preservation and protection of such practices/results to be obligatory (or morally significant independent of some social practice). In such cases, the practice is said to enter into the thinking of the agent in a merely derivative way.

One wonders whether this line of thought threatens to turn most social norms into moral norms. The concern may be put in terms of the emerging literature on the general motivations for norm conformity. It seems plausible that much norm conformity is motivated by a suite of emotions that are not clearly distinct from the kinds of moral dimensions that Haidt has discussed: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and purity/degradation. The evolution of the disposition to these may be associated with several adaptive challenges: "caring for vulnerable children, forming partnerships with non-kin to reap the benefits of reciprocity, forming coalitions to compete with other coalitions, negotiating status hierarchies, and keeping oneself and one's kin free from parasites and pathogens, which spread quickly when people live in close proximity to each other" (Jonathan Haidt, *The Righteous Mind* [New York: Random House, 2011], 112). It certainly seems plausible that the grounding motivation of many norm adherents turns on sensitivities to things like harmfulness, loyalty, authority, and fairness to those with whom one is mutually reliant. In such cases, do the norms apparently constitute moral norms? There is much to be sorted out here. But the issues and discussion are rewarding.

As befitting its title, the second part of *Explaining Norms* is focused on the dynamics by which norms arise and evolve. The authors seek to draw on elements from two different approaches to the dynamics of norms—the rational purposeful action approach and the 'social meanings' approach. Ultimately, it is said, "a full explanation of norms will inevitably have to incorporate elements of both in some measure" (91).

The authors focus on the dynamics of informal norms. To understand their dynamics, it is said, one must understand how folk, as a folk or a 'we', come to authorize themselves to issue the principles or norms and interpret and apply those principles to each other and themselves individually. These groups must get enough of the individuals that comprise them to come to accept the principle and to regard certain applications as the correct ones and to hold others accountable in ways that enough think fitting (94–95). This is characterized as "bootstrapping norms."

Norms might be bootstrapped in several ways—and the authors discuss four. First, there is a relatively formal two-step process possible: when actors first obtain a set of articulate 'I will if you will' promises, then note that enough others have so committed, so that enough of the actors have effectively accepted the rule that is the focus of the norm. This process seems most characteristic of nation-states as agents settling on norms or conventions. Second, there are cases in which norms emerge by way of "free-flowing cascades." In such cases there is

among the group a “bandwagon-style mechanism driven by differential sensitivities” (98). Here one finds sensitivities to various costs, benefits, and ranges of information—including information about others’ sensitivities and conditional acceptances (99). A third possibility involves something like a cascade among a “power-differentiated” class who then leads in arriving at the emerging norm. Finally, a fourth route to the formation of a norm involves the transformation of a social practice or convention into a norm. Here “non-normative behavioural regularities that serve a coordination-facilitation function metamorphose into something normative that serves an accountability-creating function” (101).

The authors discuss several bases for norm persistence turning on actor’s investments and sunk costs associated with extant norms, the costs of converting were other norms adopted, interests and loss aversion, and self-fulfilling expectations informed by extant norms. The book also provides a useful characterization of the ways in which norms can change or unravel. One interesting twist here involves the relation between norms of various kinds and associated social practices. While these are treated as conceptually distinct matters, it is acknowledged that norms often are conditioned by expectations regarding the behavior and thought of others. Thus, insofar as folk’s expectations are sensitive to actual patterns of social practice, changes in social practice can unravel norms (113–18).

Explaining norms is thus a matter of understanding processes of the sort just crudely categorized. One approach to understanding these processes is to attempt a rational reconstruction of them. This would “show how the norm solves a specific problem that purposive agents might have—like helping them coordinate their behavior or cooperate despite the temptation to cheat—and how it is in most people’s interest to comply with it” (133). Here, one treats the norm as a solution to a game-theoretic problem faced by a set of rational agents.

As the authors note, the explanation afforded is typically understood as an explanation of why norms emerge in light of an agent’s purposes, not of how they emerge. As such, the explanation would provide an incomplete understanding of norms and their dynamic. But, the authors also focus on something of a problem or limitation of such purposive explanations: “not all norms are unqualifiedly good ones, in that not all norms provide a real or unqualified solution to game-theoretic problems faced” (143). Ultimately, the authors suggest that the rational reconstructions are only explanatory insofar as they dovetail with understandings of how the relevant norms arise and that this turns upon identifying the mechanism by which norm acceptance emerges within the group in question. How-questions and why-questions cannot be neatly separated (143–47, 154).

A second broad approach to understanding the processes by which norms emerge and change proceeds in terms of social meanings—what the norms and practices “say about” one’s self, one’s community, one’s place in one’s community, and the like. The authors suggest that in accounting for the full dynamic of norms, one will need to make use of both approaches.

DAVID K. HENDERSON  
*University of Nebraska, Lincoln*