
CONVERSATION 9

**Commitment, Identity, and Collective Intentionality:
The Basis for Philanthropy**

Paul Lewis

COMMENTS

Don't Explain!

Kevin Quinn

What is Wrong with a Social Preference, or Two?

Shaun P. Hargreaves Heap

*Conversations on Philanthropy Volume VI:
Identity, Interests, & Philanthropic Commerce*

© *DonorsTrust* 2009

CONTRIBUTORS

SHAUN P. HARGREAVES HEAP is a graduate of Oxford and UC Berkeley. He is Professor of Economics at the University of East Anglia and has held positions at Concordia University and the University of Sydney. His current research concerns the social influences on decision making and in particular involves an experimental investigation of how group membership affects people's willingness to trust each other. He has published research in this area recently in the *Economic Journal* and in the *American Economic Review*. His other area of current research is on the influence and measurement of diversity in the media.

PAUL LEWIS was educated at Peterhouse, Cambridge, and Christ Church, Oxford. He was a Newton Trust Lecturer in the Faculty of Economics and Politics, and the Faculty of Social and Political Sciences, Cambridge University, and a Fellow of Emmanuel and Selwyn Colleges, before becoming a Senior Lecturer in Economics at King's College, University of London. His research interests include: the Austrian school of economics; applied microeconomics; and the history and methodology of economics. He is a Visiting Scholar at the Mercatus Center at George Mason University, USA, a retained supervisor in economics at Peterhouse, Cambridge, and a member of the Cambridge Social Ontology Group.

KEVIN QUINN is Associate Professor of Economics at Bowling Green State University. His research interests are in economic philosophy, especially the nature of rationality, and in the history of economic thought.

COMMITMENT, IDENTITY, AND COLLECTIVE INTENTIONALITY: THE BASIS FOR PHILANTHROPY

Paul Lewis

Philanthropy—“voluntary giving and association that serves to promote human flourishing” (Ealy 2005, 2)—is often said to have three key features. First, it is said to involve behavior that is in some sense “disinterested,” “benevolent,” or “altruistic.” For Steven Grosby, for instance, the existence of philanthropy reveals that there are “a number of diverse orientations of human action,” one of which takes the form of “disinterested or ‘selfless’ action, of works on behalf of others whom one knows or does not know which imposes costs on the individual but which confers benefits to others” (2009, 2, 4). Similarly, according to Jacques Godbout the gift is “a moral act and as such is ‘intrinsically motivated and not subject to means-ends analysis’ ” (1998, 91).

The second feature of philanthropy that is often highlighted concerns the origins of the motivation for such disinterested behavior. The latter is often said to derive at least in part from the way in which people form attachments to one another and thereby forge a sense of identity (that is, a sense of who they are). According to Godbout, for example, we should “consider the gift mainly in terms of the dimension of donation as part of self expression” (1998, xi). Kenneth Boulding elaborates on this point by noting that philanthropy is “an expression of [a person’s] sense of community with others” (1974 [1962], 240), an endeavor involving a non-calculative act of benevolence “in which the decision maker elects to do something not because of the effects the decision will have in the future but because of what he ‘is’ here and now, [that is, because of] how he perceives his own identity” (1970, 132):

It is to the subtle dynamics of the integrative system—that set of social relations involving status, identity, community, legitimacy, loyalty, and trust—that we have to look if we are to understand the growth and structure of the grants economy (1973, 5).

Similarly, Lenore Ealy observes that people’s actions are driven by a variety of motives—some benevolent, some not—and each of us has to “struggle with our conflicting drives and ends as we seek to understand and forge our own identities” (2007, viii).

According to this view, benevolence—like other forms of self-sacrificing, group-oriented action—is said to reflect “an essential part of man’s nature—his need to identify with others, his need to expand his interests and concerns beyond the sphere of his own body” (Boulding 1974 [1965], 250). It follows, and here is our third point, that understanding such behavior requires us to acknowledge an “expanded notion of the self” (Grosby 2009, 1, 5) that—in contrast to the isolated, hermetically sealed atom that is *homo economicus*—is constituted at least in part by the relations in which the individual stands to other people.

However, while (as we have seen) these three aspects of human nature are regarded as significant for understanding philanthropy, writers on the topic have yet to provide a coherent account of how these elements combine to give rise to philanthropic activity. It is this lacuna that I aim to remedy in this essay (cf. Garnett 2009, 7, 9). In what follows I shall draw on the work of a number of social theorists and philosophers, most notably the Nobel Prize-winning economist Amartya Sen, in order to outline several concepts—specifically the notions of “commitment,” “collective intentionality,” and “identity”—that will help us to elaborate on those ideas in order to provide a more coherent conceptual framework for thinking about the nature of philanthropic action. The following section of this essay offers an account of Sen’s work on the nature of rationality, focusing on his distinction between sympathy and commitment. The section after that elaborates on the notion of commitment by drawing on the philosophical literature on collective intentions to bring out the links between commitment and personal identity. It then goes on to suggest how a person’s efforts to forge their identity might involve them engaging in philanthropic behavior. In the subsequent section I attempt to allay some of the concerns that classical liberals might have about the role of philanthropy, by pointing out that—contrary to what some classical liberal authors suggest—the type of (other-regarding) motivation that underpins and gives rise to philanthropic action also plays a significant role in the generation of the orderly outcomes in decentralized market economies. The final section offers a few concluding remarks.

Varieties of Motivation: Sen on Sympathy and Commitment

Sen begins by arguing that the standard account of the nature of self-interest typically provided by economists elides three different ways in which considerations of one's own "self" enter into reasoned decision-making, each of which reflects a different degree of "privateness" (that is, a different degree to which the self alone is central to a person's behavior). The first, and narrowest, account of the role of the self rests on the assumption of *self-centered welfare*, which holds that a person's welfare depends only on his or her *own* consumption, ruling out the possibility that the person's welfare is affected either by considerations of sympathy (or antipathy) for others or by a personal concern for more abstract, distributional issues pertaining to social justice. The second conception of the self rests on the assumption of a *self-welfare goal*. The latter permits the individual to display more concern for the rest of the world than does the assumption of self-centered welfare, and so permits a broader notion of the self to figure in people's decisions, because, while stipulating that a person's only goal is to maximize his or her own welfare, it allows that the latter may be directly affected by the welfare of others. That is, it permits the possibility of *sympathy*, a notion that will be discussed at greater length below. The third assumption is that of *self-goal choice*, according to which a person's choices must be based only on the pursuit of his or her *own* goals. This assumption excludes the possibility that a person's choices may be influenced by factors other than his or her own goals, in particular by self-imposed restrictions on the pursuit of one's own goals (reflecting, for instance, the influence of social norms and rules that discourage the pursuit of goals that the individual, taken in isolation, would prefer to pursue) (Sen 2002a, 33-35).

Sen elaborates on the nature of such other-regarding behavior by drawing a distinction between sympathy and commitment (1977, 326-29). As noted above, sympathy involves one person's welfare being affected by the welfare of others—as, for example, when one gains pleasure from observing the happiness of others. Of necessity, sympathy involves a departure from the assumption of self-centered welfare, though it leaves intact the assumptions of self-welfare goal and self-goal choice. Commitment, on the other hand, denotes a person's willingness to act in a particular way not because doing so maximizes his or her welfare but simply because the type of action in question is required for conformity with a social rule or norm that the individual regards as inviolable. While behavior that is driven

by commitment need not violate the assumption of self-centered welfare, it *does* involve a breach of the assumptions of self-welfare goal and self-goal choice because it involves people refraining from pursuing the (individually chosen) goals that would maximize their personal welfare, in favor of adhering to social norms and rules. Commitment therefore “drives a wedge between personal choice and personal welfare” in the sense that it leads to a person “choosing an act that he believes will yield a lower level of personal welfare to him than an alternative that is also available to him” (329; also see 2002a, 35; 2005, 7).¹ And as we shall discuss in more detail below, by violating the assumption of self-goal choice, Sen’s approach permits a broader conception of the self than does *homo economicus*, in the sense that it allows that a person may reason from a vantage point taking into account not only their narrow self-interest but also what is best from the point of view of a *group* of people of which that individual is a member.

More specifically, the possibility of commitment is related to a fourth aspect of the self or person distinguished by Sen. In developing his account of commitment, Sen emphasizes that people are reflexive beings who are able to scrutinize themselves and in particular to reason about the propriety of their desires and values. According to this view, far from reason being the slave of the passions, people are able to use their power of reason to reflect upon, to control, and even to override their desires and individual goals, so that the latter do not always manifest themselves in people’s actions:

A person is not only an entity that can enjoy one’s own consumption, experience, and appreciate one’s welfare, and have one’s goals, but also an entity that can examine one’s values and objectives and choose in the light of those values and objectives. Our choices need not relentlessly follow our experiences of consumption or welfare, or simply translate perceived goals into action. We can ask what we want to do and how, and in that context also examine what we should want and how (Sen 2002a, 36; also see 25 and 40-42).

Indeed, according to Sen (2002a, 5-7, 50-52) it is precisely this capacity to reflect upon one’s values and actions and to make commitments to shared social rules and norms that makes someone a person (Frankfurt 1971). By ignoring this fourth aspect of the self, and therefore in effect denying the possibility that people can reason about their desires and goals, standard rational choice theory “involve[s], in effect, a basic denial of *freedom of thought*” (Sen 2002a, 5):

[T]he insistence on the pursuit of self-interest as an inescapable necessity for rationality subverts the “self” as a free, reasoning being, by overlooking the freedom to reason about what one should pursue (46; also see 37).

For Sen, therefore, we can do justice to the influence of reason on human action only if we acknowledge that the domain of reason extends beyond the (instrumental) task of assessing the most effective means of satisfying preferences and goals that are not themselves given by reason to embrace the possibility of (deontological) reasoning *about* the values and goals that people (should) choose to pursue (2002a, 40; 2006, 20-22; also see Hirschman 1984).

In explaining why people might depart from the instrumentally rational mode of conduct presupposed by rational choice theory, Sen emphasizes that, far from being isolated atoms, people are social beings whose values, goals, beliefs, and actions are all shaped by the network of social relations within which they are embedded (cf. Godbout 1998, 16-18). More specifically, according to Sen one important reason why a person might refrain from pursuing her individual (self-)goals in favor of conforming to social rules and norms is because doing so enables her to form attachments to various groups and thereby to cultivate her identity (that is, a sense of who she is) (2002 [1985], 215; also see 211-12; 2002a, 40-41; 2005, 5, 7). The connection between a person’s commitment to social rules, on the one hand, and his or her identity, on the other, has been fruitfully analyzed by philosophers working on the notion of collective intentionality, to which concept we now turn our attention.

Collective Intentionality, Identity, and Commitment

The philosophical literature on collective intentionality suggests that in addition to having the individual intentions—the individual commitments to (purposive) action—that economists typically attribute to them, the individual members of a group might also possess “shared” or “collective” intentions that involve them making a commitment to act in concert with one another, as a group (Davis 2002; Gilbert 1989, 1996). The distinction between individual and collective intentionality often arises in everyday life: people use the language of *individual* intentionality when they view themselves as acting independently of others (“I want,” “I hope,” etc.), and they invoke collective intentionality when they see themselves as acting as part of a group (“what we want,” “what the

Austrian school aspires to do,” etc.). As we are about to see, the theory of collective intentionality suggests that, far from being an insignificant linguistic trifle, the distinction between “I” and “we” intentions signifies an important difference in the type of motivation driving people’s actions, the recognition of which can enable us to conceptualize the roles of commitment, identity, and (a broader notion of) the self in reasoned decision-making.

Collective intentions—or *we-intentions* as they are also known—have two key characteristics. First, an individual who expresses a we-intention believes that the intention in question is widely, though not necessarily universally, held by the other members of the group. Second, the individual believes that the intention is *mutually* or reciprocally held by members of the group, in the sense that they too believe that it is widely held by their fellow members. Thus, for example, if Ludwig, who is a member of a group of Austrian economists, says, “We believe in the efficacy of the free market,” Ludwig is saying that he believes in the efficacy of the free market and that, in addition, he believes that the other members of the group both believe in the efficacy of the market themselves and also attribute that belief to one another (cf. Davis 2002, 14). In a nutshell, we-intentions involve a structure of mutually reinforcing, reciprocal beliefs shared by the individual members of the relevant group, such that each believes that the others hold the same belief and each also believes that the others think the same about their fellow members (cf. Shils 2006, 197-98).

Significantly, and in keeping with Sen’s notion of commitment, behavior driven by we-intentions is not reducible to instrumentally rational behavior (Searle 1995, 23-25; Davis 2002, 20-22; 2004, 392-93). The reason is as follows. An individual’s we-intention centers on what (s)he thinks the intentions of the other individuals in the group actually *are*, not what (s)he would *like* them to be, so that there arises the possibility of a tension between what an individual *believes* a group’s collective intention to be and what (s)he would *prefer* it to be. If an individual (sincerely) expresses a we-intention in a situation where that tension has indeed arisen, then (s)he has effectively made a (Senian) commitment to act in accordance with the group’s collectively expressed view that a particular goal should be pursued, or that a particular type of action is required, and so forth, even though that might not be what the individual would have preferred were (s)he not a member of the group in question. According to this view, the shared intentions that arise when a person uses we-language

involve him imposing upon himself obligations or commitments that qualify the unconstrained pursuit of his own (self-)goals, simply because expressing a we-intention requires an individual to conform to how other people use that same “we.” Hence, people may share a collective goal without each of them also having it as a personal goal (Davis 2002, 21-22; 2004, 399; Anderson 2003, 191-93).

The notion of collective intentionality is significant in understanding the role of identity and commitment in human action because one way of conceptualizing a person’s identity is in terms of the social groups with which she chooses to affiliate herself. Such affiliations can in turn be thought of as involving the use of first-person plural speech in order to form collective intentions about what “we” want, believe, and so forth (Davis 2007). As Sen has put it,

The nature of our language often underlines the force of our wider identity. “We” demand things; “our” actions reflect “our” concerns; “we” protest at injustice done to “us.” This is, of course, the language of social intercourse and politics, but it is difficult to believe that it represents nothing other than a verbal form, and in particular no sense of identity (2002 [1985], 215; also see 2002a, 41).

According to this view, people’s capacity to identify themselves with others, and thereby to define who they are, is captured by the way their use of we-language requires them to embrace—in Sen’s sense of commitment—the intentions of the other group members (to whom the “we” is meant to apply) about what goals to pursue, what counts as acceptable behavior, and the like. For the (sincere) use of such language requires a person to adopt the same (joint) standpoint as others in the group, “accept[ing] as reasons for action only those considerations that each person would be willing to accept as reasons for everyone to act” (Anderson 2001, 29). We-intentions can thus be seen to involve people transcending the narrow confines of their own self-interest in order to consider what is right or best not from their own point of view but from the perspective of the group as a whole. In Sen’s words, “the pursuit of private goals may well be compromised by the consideration of the goals of others in the group with whom the person has some sense of identity” (2002 [1985], 215; also see 214 and 2002a, 40-42).

It is especially noteworthy in this regard that a person’s membership in social groups is often conditional upon his or her faithfully observing various social rules and norms (Sen 2002 [1985], 216-17). The latter can be expressed in

terms of collective intentions to the effect that “we believe that members of the group should do x in circumstances z” (where, in addition to being read literally, the phrase “do x” should be interpreted broadly as a placeholder for a variety of injunctions such as, “count as,” “take to mean,” “refrain from,” “donate to,” and so forth) (Davis 2004, 390; cf. Lawson 2003, 36-39). Such rules and norms that specify what types of behavior the members of a particular group count as “correct,” “honorable,” “just,” etc., and in effect constitute a set of guidelines or a script that tells people what they have to do in order to identify themselves with the (other members of the) group and thereby to cultivate and express publicly their identity as group members (Sen 2002 [1985], 215). Such rules and norms have motivational force because they furnish people who wish to become or remain members of a particular group with reasons for acting in certain ways:

One of the ways in which the sense of identity can operate is through making members of a community accept certain rules of conduct as part of obligatory behavior towards others in a community. It is not a matter of asking each time, What do I get out of it? How are my own goals furthered in this way?, but of taking for granted the case for certain patterns of behavior towards others (216-17).

More specifically, the motivational force of social norms derives from the fact that group members accept the authority of “us”—of “our” shared view of how “we” should behave—to determine (key features of) their conduct in the domain defined by the norm. As Elizabeth Anderson has put it, “To count as a reason for action, a consideration must appeal to a person’s self-understanding, not [necessarily] her self-interest. It must fit into her understanding of her identity” (2003, 192; also see 193).

In this way the notion of collective intentionality makes it possible to conceptualize how people can have sources of motivation—including, as we shall see, those that enjoin them to engage in philanthropy—above and beyond the instrumental desire to satisfy their preferences. People not only have the capacity to behave in an instrumentally rational fashion, asking what should *I* do and striving to satisfy their own preferences, they also have the (often countervailing) ability to act in accordance with social rules and collective goals, their commitment to which may involve stepping back from their individual goals and asking what is the best strategy for *us* to adopt:

Behavior is ultimately a social matter as well [as an individual one], and thinking in terms of what “we” should do, or what should be “our” strategy, may reflect a sense of identity involving recognition of other people’s goals and the mutual interdependencies involved (Sen 1987, 85; also see 2002 [1985], 41).

The members of a social group think of themselves as a “we” and understand one another to be jointly committed to various goals, including that of upholding shared social rules and norms. In identifying with a group, therefore, an individual understands that she is accepting responsibility for doing her part to advance the group’s goals and to uphold the rules and norms that operate within it, making a commitment that motivates her subsequent actions. Hence, as Sen has put it, “the sense of identity takes the form of partly disconnecting a person’s choice of actions from the pursuit of self-goal” (2002 [1985], 216).

Overall, then, the vantage point provided by the theory of collective intentionality suggests that Sen can be thought of as advocating a relational theory of (social) identity whereby, through their use of we-language, people express their joint commitment to various goals and social rules and norms, forging links to the groups in which those rules and norms prevail and thereby conferring identity upon themselves. To put this point slightly differently, Sen is arguing that the possibility of identifying with others presupposes an extended notion of the self that, in contrast to the isolated atom that is *homo economicus*, is forged at least in part by a person’s affiliations with various social groups (cf. Grosby 2009, 1, 5). Moreover, as we have seen, thanks to her attachment to such groups, the (socially conditioned) person postulated by Sen is able (to an extent, at least) to adopt a more disinterested vantage point that makes it possible for her to act in accordance not only with her narrow self-interest but also with what is right or best from the point of view of a group—or indeed society—as a whole, treating other people not merely as (intrinsically worthless) instruments for the realization of her own ends but as ends in themselves whose own goals have intrinsic value. As Grosby (3) has put it, “the attachments or social relations constitutive of some groups indicate an orientation of human action beyond that of the interest of the individual (in this sense, transcending the self) ... such that considerations of what is right or of a common good can co-exist with individual self-interest” (see Godbout 1998, 7-9, 13, 20). Viewed thus, the

notion of we-intentions can be thought of as one way of conceptualizing the “imaginative capacity” of the mind that, according to commentators such as Boulding (1974 [1962], 239-40), Shils (2006, 213), and Grosby (2009, 2-3), enables us to identify ourselves with others and thus gives rise to the possibility of behavior in their service, motivated by ideals of charity, fairness, *noblesse oblige*, and so forth (rather than by the prospect of personal gain).

In particular, according to Kenneth Boulding it is just such a sense of identity that is one of the wellsprings of philanthropy. “A gift helps to create the identity of the giver, and a gift either to an individual or to a cause or to a community identifies the giver with the recipient.... Thus, the gift builds itself into the identity of the giver” (1973, 27-28). For example, the members of a particular social group—the alumni of a college—might think of themselves as a “we” and understand one another to be jointly committed to various goals, including that of supporting their college through benefactions. As Grosby has put it, in language redolent of the theory of collective intentionality, philanthropy “points to attachments formed by individuals [that are expressed in the form of statements involving] ... a ‘we’, implying an expansion of the individual’s conception of the self” (2009, 2, 3; also see Godbout 1998, 16). Whereas self-interested, utility-maximizing behavior involves a person effectively identifying with him- or herself, commitment involves a person identifying with others. In identifying with their college and with their fellow alumni, an individual understands that she has accepted responsibility for doing her part to advance the group’s goals, in this case by upholding the group norm that stipulates that members should give back to their *alma mater*. Giving a gift to their old college, for example, is one way for alumni to reaffirm their affiliation for it, thereby cultivating their identity as members of that institution. In this way, as Boulding has put it, giving is “a sacrifice we may make in the interests of our identity, for our identity depends very largely on the community with which we identify” (quoted by Ealy 2007, viii).

Moreover, if Boulding (1973, 26-27) is right in arguing that the “we” invoked by alumni when they engage in such altruistic joint giving embraces not only a person’s immediate contemporaries but also other generations of members, then adhering to the norm also enables one to identify with earlier generations (from whose beneficence one benefited) and also with future generations (who will benefit from one’s generosity), thereby becoming a

member of a community that endures over time. In this way, giving to one's college exemplifies Boulding's point that "in order to establish a satisfactory identity, one must maintain some sort of community, however uncertain and discounted, not only with one's own day, but with the whole human race as it stretches out through time and space" (97). In this way the notion of collective intentionality might be thought of as one way of conceptualizing the "integrative relationships[,] ... that is, [the] groups of people who have some feelings of identification and benevolence towards each other," which, according to Boulding, lie at the heart of philanthropy and gift-giving (1973, 27; cf. Godbout 1998, 69-72, 94).

Commitment and the Extended Order

In this final section I want briefly to consider a potential objection that classical liberals might advance to the account of philanthropy outlined above, namely that the type of (other-regarding) motivation which (it has been suggested) underpins philanthropy is relevant only in those cases where a small number of people, who are well known to one another, are pursuing an agreed-upon goal (Hayek 1988). In particular, Hayek posits that as soon as we move from the realm of such face-to-face interactions to the impersonal exchanges that characterize the extended order of the Great Society, altruistic impulses of the sort considered above can no longer be relied on, not least because of the difficulty of acquiring the information required to dispense one's largesse efficiently, and must be replaced by an emphasis on the pursuit of narrowly defined self-interest.

The point I want to make here is that despite the account of the shortcomings of altruism found in Hayek's explicit remarks on philanthropy, his own account of how the extended order of the Great Society is possible reveals a more balanced position, in which other-regarding forms of motivation *do* have a significant role to play. Significantly, Hayek does *not* contend that the pursuit of self-interest should be given free reign, even in the trades between anonymous others that characterize decentralized market economies. On the contrary, and in keeping with Sen's point (2005, 7; 2007, 347-52) that in living in society individuals come to recognize that there are other people who are trying to pursue their own goals and that they should be given a fair opportunity to do so, a *desideratum* that requires one to refrain from the exclusive pursuit of

one's own (self-)goals, Hayek acknowledges that the generation of an orderly allocation of resources in decentralized market economies requires that people's pursuit of their goals be tempered by their willingness to abide by—one might even say, feel a commitment to—certain abstract norms and rules (moral norms such as promise-keeping and truth-telling; generalized norms of reciprocity; and the laws of property, tort, and contract) (Hayek 1960, 62-63; 1976, 14, 16-17; also see McCann 2002, 11-19; and Lewis and Chamlee-Wright 2008, 110-112).

Indeed, as Paul Seabright (2004, 56-58) has argued, no economy can function by relying on self-interest alone, simply because the successful implementation of trades ultimately depends on someone, somewhere—whether it be the trading parties themselves or officers of the legal system—feeling under a moral obligation to honor and/or enforce the terms of the contract. As Kenneth Arrow has put it, “there has to be some kind of commercial morality for contracts to be executed[;] ... a theory which depends merely on reputation is not enough because there will always be circumstances when it pays to violate the rule.... So the economic system—the self-seeking, *laissez faire* system—would not work without the presence of these non-*laissez faire*, non-self-seeking norms” (1990, 139; also see Dobuzinskis 2009, 132-133).²

Moreover, Hayek appears to conceptualize the way in which rules influence people's actions in a way similar to that used by Sen. In *Law, Legislation, and Liberty*, for example, Hayek elaborates on his comment that “Man is as much a rule-following animal as a purpose-seeking one” by quoting the following passage from the philosopher R. S. Peters:

Man is a rule-following animal. His actions are not simply directed towards ends; they also conform to social standards and conventions, and unlike a calculating machine he acts because of his knowledge of rules and objectives. For instance, we ascribe to people traits of character like honesty, punctuality, considerateness and meanness. Such terms do not, like ambition, or hunger, or social desire, *indicate the sort of goals that a man tends to pursue; rather they indicate the type of regulations that he imposes on his conduct whatever his goals may be* (1973, 11, 147-48 n. 7; emphasis added).

For Hayek, then, social rules and norms are not goals that people pursue, but rather regulations that people impose upon themselves in order to restrain

the ardor with which they pursue their own goals:

The rules of morals are instrumental in the sense that they assist mainly in the achievement of other human values; however, since we only rarely can know what depends on their being followed in the particular instance, to observe them must be regarded as *a value in itself, a sort of intermediate end which we must pursue* without questioning its justification in the particular case (1960, 67, emphasis added; also see 1976, 16-17).

Such passages suggest, therefore, that it is not wholly unreasonable to view the social rules and norms that Hayek mentions in his discussion of the extended order as examples of the type of self-imposed commitment to rules and norms discussed by Sen, a commitment that involves a violation of the assumption of self-goal choice.

It is also worth noting in this regard that, according to Sen, a person can follow social rules and norms that require him to curtail the vigor with which he pursues his own goals, thereby making it easier for others to pursue *their* goals without that first individual taking the goals of those other people as his own. On the contrary, according to Sen,

One can take note of other people's goals and priorities and decide to constrain the unifocal pursuit of one's own goals with behavioral constraints and other restrictions, without that self-restraint being interpreted as the pursuit of the goals of others.... Rather, you are just following a norm of good behavior you happen to approve of (to wit, "let others be"), which is a self-imposed restraint you end up accepting in your choice of what to do (2007, 347-49).

For Sen, then, adhering to social rules and norms that enable others to pursue their goals more effectively does not mean that one is taking their goals as one's own. Instead, "the violation of self-goal choice is arising here from the normative restraint we may voluntarily impose on ourselves on grounds of recognizing other people's pursuits and goals, without in any substantive sense making them our own goals" (2007, 353-54; also see 2002 [1985], 214). Again, this interpretation seems quite consistent with Hayek's point that because the codes of conduct that govern people's behavior in the extended order are abstract in the sense that they do not require people to agree on the (concrete) ends which they are pursuing, they enable people to pursue diverse goals in

peace and harmony (1973, 1988). Like Sen, Hayek is arguing that people are more likely to be able to pursue more of their own goals successfully if they voluntarily submit to the restrictions imposed on self-goal choice by certain social norms and rules.

It is perhaps noteworthy in this regard that in his account of the importance of rule-governed behavior Hayek defines rationality in a way that does not reduce it to the instrumental version that characterizes *homo economicus*: “Rationality ... can mean no more than some degree of coherence and consistency in a person’s action, some lasting influence of knowledge or insight which, once acquired, will affect his action at a later date and in different circumstances” (1960, 77). This definition is not so far removed from the one provided in Sen’s account of rationality as involving the discipline of subjecting one’s choices—of goals and values, as well of actions—to reasoned scrutiny. As Sen writes of Adam Smith, “rationality is seen as reasoned reflection on the nature of the processes involved and the consequences generated, in the light of the valuations one has reason to accept, ... not a fixed formula with a pre-specified maximand” (Rothschild and Sen 2006, 258).

Hayek’s own account of the possibility of social order in decentralized market economies reveals, then, that the binary opposition he draws, in his explicit comments on altruism and philanthropy, between other-regarded motives (whose role he contends ought to be confined to the micro-cosmos of face-to-face interaction between personal acquaintances) and self-interest (the pursuit of which ought to be people’s sole concern in the extended order) is in fact overdrawn. This opposition needs to be qualified because, as we have seen, Hayek’s own account of the extended order arguably allows—indeed, requires—that there be constraints on the pursuit of self-interest similar to those involved in philanthropic activity. And if other-regarding motivations do indeed have a role to play in a Hayekian account of the generation of social order in decentralized market economies, then the view—which Hayek at times advances—that the commercial (for-profit) and philanthropic (not-for-profit) spheres are mutually exclusive is surely overdrawn (cf. Garnett 2009, 4-7).³

Conclusion

Whether at any given moment in time a person’s behavior is influenced more by their (deontologically rational) commitment to prevailing social norms

and rules or by their (instrumentally rational) desire to satisfy their own preferences, cannot be established *a priori* but instead must be established *ex posteriori* according to the context in which people are acting (Anderson 2001, 30-31; Sen 2002a, 25-26, 47; 2006, 22-23; Godbout 1998, x). But by providing conceptual space for a broader notion of the self—one that is permeable to external influences and because of that is able to reason about the propriety of their goals and values—Sen allows for the possibility of disinterested behavior of the kind that can be argued to underpin philanthropy. The key question that remains concerns identifying the institutional context that is most conducive to such behavior both in the sense that it enables and encourages people to act upon their altruistic impulses and that it provides the information required for people to do so in an informed way. But that is a question for another time.

NOTES

- ¹ As Sen notes (1997, 760), his distinction between sympathy and commitment parallels that drawn by Adam Smith (1975 [1790], 191) between sympathy (understood as ultimately self-interested benevolence) and “generosity,” where doing things for others may involve us “sacrific[ing] some great and important interest of our own.”
- ² Kenneth Boulding (1973, 28) makes a similar point when he contends that “without the kind of commitment or identity which emerges from sacrifice [in the sense of being willing to enforce or abide by a social norm even if it is not in one’s interests to do so], it may well be that no communities, not even the family, would really stay together.”
- ³ For more on the possibility of feedback mechanisms that might help to inform and so guide philanthropic activity, see Chamlee-Wright and Myers (2008).

REFERENCES

- Anderson, Elizabeth. 2001. “Unstrapping the Straightjacket of ‘Preference’: A Comment on Amartya Sen’s Contributions to Philosophy and Economics.” *Economics and Philosophy*, 17 (1): 21-38.
- . 2003. “Beyond Homo Economicus: New Developments in Theories of Social Norms.” *Philosophy and Public Affairs*, 29 (2): 170-200.
- Arrow, Kenneth J. 1990. “Kenneth J. Arrow.” In R. Swedberg, *Economics and*

- Sociology*. Princeton, NJ.: Princeton University Press.
- Boettke, Peter J. and Virgil H. Storr. 2002. "Post-Classical Political Economy: Polity, Society and Economy in Weber, Mises and Hayek." *American Journal of Economics and Sociology*, 61 (1): 161-91.
- Boulding, Kenneth. 1970. *Economics as a Science*. New York: McGraw-Hill.
- . 1973. *The Economy of Love and Fear: A Preface to Grants Economics*. Belmont, Cal.: Wadsworth Publishing Company.
- . 1974 [1962]. "Notes on a Theory of Philanthropy." In F. R. Glahe, ed., *Collected Papers of Kenneth Boulding, Volume II*. Boulder, Co.: Colorado Associated University Press.
- . 1974 [1965]. "The Difficult Art of Doing Good." In L. D. Singell, ed., *Collected Papers of Kenneth Boulding, Volume IV*. Boulder, Co.: Colorado Associated University Press.
- Chamlee-Wright, Emily and Justus A. Myers. 2008. "Discovery and Social Learning in Non-Priced Environments: An Austrian View of Social Network Theory." *Review of Austrian Economics*, 21 (2-3): 151-66.
- Cornuelle, Richard C. 1993 [1965]. *Reclaiming the American Dream: The Role of Private Individuals and Voluntary Associations*. New Brunswick, NJ.: Transaction Publishers.
- Davis, John B. 2002. "Collective Intentionality and Individual Behavior." In Edward Fullbrook, ed., *Intersubjectivity in Economics: Agents and Structures*. London and New York: Routledge.
- . 2004. "Collective Intentionality, Complex Economic Behavior, and Valuation." In Davis *et al.*, eds., *The Elgar Companion to Economics and Philosophy*. Cheltenham: Edward Elgar.
- . 2007. "Identity and Commitment: Sen's Fourth Aspect of the Self." In Fabienne Peter and H. B. Schmid, eds., *Rationality and Commitment*. Oxford: Oxford University Press.
- Dobuzinskis, Laurent. 2009. "French Perspectives on the Origin and Logic of the 'Philanthropic Order': A Critical Account." *Conversations on Philanthropy VI*: 115-139. ©2009 DonorsTrust.
- Ealy, Lenore T. 2005. "The Philanthropic Enterprise: Reassessing the Means and Ends of Philanthropy." *Economic Affairs* 25 (2): 2-4.

- _____. 2007. "Introduction." *Conversations on Philanthropy IV*: v-ix. ©2007 DonorsTrust.
- Frankfurt, Harry. 1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy*, 68 (1): 5-20.
- Garnett, Robert. 2009. "Philanthropy, Markets, and Commercial Society: Beyond the Hayekian Impasse." (unpublished paper, Department of Economics, Texas Christian University.)
- Gilbert, Margaret. 1989. *On Social Facts*. London: Routledge.
- _____. 1996. *Living Together: Rationality, Sociality, and Obligation*. Lanham, MD: Rowman and Littlefield.
- Godbout, Jacques T. 1998. *The World of the Gift*. Montreal & Kingston: McGill-Queens University Press.
- Grosby, Steven. 2009. "Philanthropy and Human Action." *Conversations on Philanthropy VI*: 1-14. ©2009 DonorsTrust.
- Hargreaves Heap, Shaun. 2000. "How Far Can You Get with Hermeneutics?" In Peter E. Earl and S. F. Frowen, eds., *Economics as an Art of Thought*. London and New York: Routledge.
- _____. 2001. "Expressive Rationality: Is Self-Worth Just Another Kind of Preference?" In Uskali Mäki, ed., *The Economic World View: Studies in the Ontology of Economics*. Cambridge: Cambridge University Press.
- _____. 2002. "The Reality of Common Cultures." In Uskali Mäki, ed., *Fact and Fiction in Economics: Models, Realism and Social Construction*. Cambridge: Cambridge University Press.
- _____. 2004. "Economic Rationality." In Davis *et al.*, eds., *The Elgar Companion to Economics and Philosophy*. Cheltenham: Edward Elgar.
- Hayek, F. A. 1948 [1945]. "Individualism: True and False." In F. A. Hayek, *Individualism and Economic Order*. Chicago: University of Chicago Press.
- _____. 1960. *The Constitution of Liberty*. London and New York: Routledge.
- _____. 1973. *Law, Legislation, and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy. Volume I: Rules and Order*. London and New York: Routledge.
- _____. 1976. *Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy. Vol. II, The Mirage of Social Justice*. London and New York: Routledge.
- _____. 1988. *The Fatal Conceit: The Errors of Socialism*. W. W. Bartley III, ed. London: Routledge.

- Hirschman, Albert O. 1984. "Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse." *American Economic Review*, 74 (2): 89-96.
- Lawson, Tony. 2003. *Reclaiming Reality*. London and New York: Routledge.
- Lewis, Paul and Emily Chamlee-Wright. 2008. "Social Embeddedness, Social Capital and the Market Process: An Introduction to the Special Issue on Austrian Economics, Economic Sociology and Social Capital." *Review of Austrian Economics*, 21 (2-3): 107-18.
- McCann, Charles 2002. "F. A. Hayek: The Liberal as Communitarian." *Review of Austrian Economics*, 15 (1): 5-34.
- Rothschild, Emma and Amartya K. Sen. 2006. "Adam Smith's Economics." In Knud Haakonssen, ed., *The Cambridge Companion to Smith*. Cambridge: Cambridge University Press.
- Seabright, Paul (2004). *The Company of Strangers: A Natural History of Economic Life*. Oxford and Princeton: Princeton University Press.
- Searle, John R. 1995. *The Construction of Social Reality*. London: Penguin.
- Sen, Amartya K. 1977. "Rational Fools." *Philosophy and Public Affairs*, 6 (4): 317-44.
- . 1987. *On Ethics and Economics*. Oxford: Basil Blackwell.
- . 1997. "Maximization and the Act of Choice." *Econometrica*, 65 (4): 745-779.
- . 2002 [1985]. "Goals, Commitment and Identity." *Journal of Law, Economics and Organization*, 1 (2): 341-55. Reprinted in Amartya K. Sen, *Rationality and Freedom*. Cambridge, Mass.: Harvard University Press.
- . 2002a. "Introduction: Rationality and Freedom." *Rationality and Freedom*. Cambridge, MA: Harvard University Press.
- . 2002b. *Rationality and Freedom*. Cambridge, MA: Harvard University Press.
- . 2005. "Why Exactly Is Commitment Important for Rationality?" *Economics and Philosophy*, 21 (1): 5-14.
- . 2006. *Identity and Violence: The Illusion of Destiny*. London: Penguin Books.
- . 2007. "Rational Choice: Discipline, Brand Name, and Substance." In Fabienne Peter and H. B. Schmid, eds., *Rationality and Commitment*. Oxford: Oxford University Press.
- Shils, Edward. 2006. "Collective Self-Consciousness and Rational Choice." In Edward Shils, *A Fragment of a Sociological Autobiography*. S. E. Grosby, ed. New Brunswick: Transaction Publishers.
- Smith, Adam. 1975 [1790]. *The Theory of Moral Sentiments*. D. D. Raphael and A. L. Macfie, eds. Oxford: Clarendon Press.

COMMENT

DON'T EXPLAIN¹

Kevin Quinn

*“[Dostoevsky’s] concern was always with what it is to be a human being—that is, how to be an actual **person**, someone whose life is informed by values and principles, instead of just an especially shrewd kind of self-preserving animal.”*

—David Foster Wallace, *Consider the Lobster*

Paul Lewis has written a very interesting and wide-ranging paper. One of its concerns, and the one I will focus on here, is to enrich and flesh out Sen’s notion of commitment by using insights from recent philosophical work on collective intentionality and identity. The context, of course, is the application to philanthropic action, which Lewis sees, plausibly, as exemplary of Senian commitment.

The pithiest characterization Sen gives of commitment—one that might have been calculated to maximally raise the hackles of the garden variety economist—is that it consists of “counter-preferential choice” (Sen 1977). He thereby disables the first line of defense against assaults, in the name of the possibility of genuinely disinterested action, on the citadel of rational choice theory: the defense that consists of expanding agents’ objective functions to allow sympathetic motives, allowing them to derive utility not just from their own consumption but also from the utility of others. Commitment, says Sen, is different from “sympathy,” where the latter term stands for action motivated by such extended preferences. It amounts to choosing what is not preferred. To defuse Sen’s challenge, it behooves a proponent of rational choice theory to explain committed choices in the theory’s own terms, where an explanation would amount to taking what looks like disinterested action and arguing that it is, after all, interested. When I discuss Sen’s article, “Rational Fools,” with my students, I find that this is something that comes naturally to them, this debunking of examples of apparent commitment along the lines of Ben Franklin’s “honesty is the best policy.”

They grasp instinctively the cynical logic of repeated Prisoner's Dilemmas and reputation building that is the stuff of what I will call *economic alchemy*: the longstanding series of attempts to turn the base lead of self-interest into the gold of cooperation.

Thus we repeat a prisoner's dilemma (PD)—infinitely. There are many equilibria in which the players cooperate—as well as one in which they always defect. In the cooperative equilibria, cooperation is enforced by the threat of future noncooperation. Given people who are sufficiently patient—who do not discount future payoffs too greatly because they are future, not present—there are all kinds of strategy pairs that yield Nash equilibria in which cooperation occurs in every period. Whether or not you find this persuasive, the requirement of infinite repetition would seem to give it only limited application. It turns out that we can get cooperation in finitely repeated games, too, but only if each of our narrowly selfish agents places some prior probability—it need not be a large number if the game is repeated a sufficient number of times (the required probability falls rapidly with the number of repetitions)—that the opponent is an unconditional cooperator. Here the tables are turned on the alchemical strategy: in a world populated exclusively by narrowly self-interested agents, and known to be such, there could never be cooperation in finitely repeated prisoner's dilemmas.

Now we might not be invested in the truth of rational choice theory, and thus not be tempted by such debunking explanations of (apparent) commitment, and yet still think that some kind of explanation is in order. The trouble is, due in part surely to the baleful influence of economists, nothing will *count as* an explanation of choice in the social sciences that doesn't in some fashion—not necessarily as crudely as in the alchemist's account—find an interest lurking behind apparent disinterest.

In other words, given what explanation has come to mean, I wonder whether we can explain the phenomenon of commitment without explaining it away. For example, here is an explanation that doesn't explain away, but that may not fit the canons of scientific explanation—and so much the worse for them, if so. Sometimes we do what is morally called for *because* it is morally called for. Of course we often deceive ourselves, convincing ourselves that what we in fact do because it is expedient is done for disinterested reasons, “because it is the right thing to do,” but a world where such a self-description would always be mistaken is not a world I would want to inhabit, and surely is not the one we do inhabit.

In light of all this, to what extent do explanations of commitment that appeal to collective intentionality and identity preserve the phenomenon to be explained?

First, as a kind of baseline, I want to take a route that Lewis doesn't take explicitly but that may be the most obvious way in which collective agency can "explain" commitment and one which most obviously fails to preserve the phenomenon. Let's take a non-repeated PD. The game is set up so that the choice is either to cooperate or defect, and rational agents will defect. The payoffs are such that defecting is better for each no matter what the other does; but each is better off when both cooperate than when both defect. If the agents are able to form a plural subject, in Gilbert's terms, they are able each to deliberate from the standpoint of the group, and they will then form the collective intention to cooperate. We would each then be acting counter-preferentially (relative to our individual preferences). But note that *We* are not acting counter-preferentially. *We* choose precisely what it is in *our* interest. So here disinterested action at the level of the individual is a function of interested action on the part of the group.²

In Elizabeth Anderson's contribution to the symposium on Sen in *Economics and Philosophy* (2001), which Lewis cites, she considers the way in which cooperation in a Prisoner's Dilemma can come about as a result of collective agency, and thus the way in which the latter can underpin Senian commitment. However, her characterization of collective agency is importantly different from what I have just given above: it is "thicker" in ways which would seem to make it immune to the charge of explaining by explaining away. "The argument," she says, "does *not* turn on the members of the group having a common aim prior to deliberation. It turns on the fact that they *accept as reasons for action only those considerations that each person would be willing to accept as reasons for everyone to act*" (29, emphasis added). The problem with this is that it succeeds in not being an "explaining-away explanation" only because the thicker notion of collective agency employed is itself predicated on Senian commitment: the individual who participates in collective agency in the sense given in the italicized passage must already be able to act counter-preferentially. Collective agency in this sense is an example of Senian commitment, not an explanation of it. What makes collective intentionality seem to be able to explain commitment without explaining it away is the equivocation on the thick and thin senses of collective intentionality: the thin ("common aim") explains, but only by explaining away, whereas the thick (Anderson's) doesn't explain away, but doesn't explain, either.

A similar dialectic arises with Lewis's suggestion that committed action stems *inter alia* from our striving to identify with a group, membership in which is partly constituted by adherence to certain non-instrumental norms. The most straightforward reading of this makes it a classic example of explaining apparent disinterest by identifying an interest. We have an interest in being part of this group, and this more encompassing interest underlies the more narrowly disinterested adherence to the norms in question. Lewis obviously intends a deeper sense of identity than this, and I think he is right. To do what he wants it to do, however, and to be true to the phenomenon, I think it will turn out that identity with the group is an expression of one's commitment to the non-instrumental principles that define the group. In that case, as with Anderson, the commitment is the *explanandum*, not the *explanans*: the identity and the collective intent follow from the commitment. For example, my commitment to the norms of scientific inquiry is bound up with my identification with the community of scientists, but the latter is just as plausibly grounded on the former as the reverse: I identify with those who share my commitment to the pursuit of truth. I certainly don't identify with those who call themselves scientists but who breach these norms. The community I identify with is constituted by the shared commitment to these norms; the identity is parasitic on the commitment.

Identity, like collective agency, has thick and thin characterizations. The thickening comes from building in normativity, which is why identity in this thicker sense can't explain commitment: it presupposes it. In addition, with thin versions of identity, besides the explaining-away problem already noted, there is the danger of explaining not commitment but sympathy. Some of Boulding's formulations on philanthropy cited by Lewis seem to exemplify this.

I gave an example above of an explanation of commitment which doesn't explain away, but I said it probably wouldn't pass muster as a "scientific" explanation. And there, indeed, is the rub. Jean Hampton (1992, 1998) and Charles Larmore (2008) have each argued that we are crippled in understanding normativity insofar as we accept a metaphysical naturalism which can make no sense of the *objectivity* of reasons. Naturalist explanations of what we have reason to do, of what we ought to do, can only make sense of what Kant called hypothetical imperatives: if you desire the end *x*, then you ought to do *y*. The authority of the norm is contingent on the existence of the desire. But what could possibly explain the authority of a categorical imperative, such as "Do *x*,

whatever your desires happen to be?” What fact about the world naturalistically conceived could establish that I ought to do x in a categorical sense? Kant himself thought that he could ground the authority of categorical imperatives on the nature of human reason itself, that we act on “laws that we give to ourselves.” The thick formulations of identity and collective agency that Lewis’s paper uses are variations on this Kantian theme.³ I am persuaded by Larmore that the Kantian project cannot make sense of the fact that, as he puts it, “reason just is our *responsiveness* to reasons” (135), that we *find* that we have reasons to x and are thereby motivated to x .

Only if we are willing, with Hampton and Larmore, to be realists about reasons, to allow that the world is not limited to matter in motion, will we be able to explain commitment without explaining it away.

NOTES

- ¹ This is the title of a song recorded by Billie Holiday. The narrator is addressing her lover, who has been cheating on her. When he starts to make up a story about where he has been, she responds, “Hush now, don’t explain./You’re my joy and pain.” She knows the “explanation” won’t explain at all.
- ² Besides being infected with interest, this way of proceeding seems to invite a regress argument. We make individual disinterest a function of group interest. But surely the group, just like the individual, may sometimes act in committed ways, counter-preferentially. Then we would seem to need a more encompassing group, the interest of which can make sense of the first group’s disinterest. Perhaps then the group of groups, the Kingdom of Humanity, can stop the regress? But even here, can’t we humans act counter-preferentially, for example, taking measures to preserve an endangered species that reduce our welfare?
- ³ See Larmore’s (2008, 112-122) critique of Christine Korsgaard’s very influential Kantian project of basing reasons on “practical identities.”

REFERENCES

- Anderson, Elizabeth. 2001. "Unstrapping the Straitjacket of 'Preference': A Comment on Amartya Sen's Contributions to Philosophy and Economics." *Economics and Philosophy*, 17: 21-38.
- Hampton, Jean. 1992. "Rethinking Reason." *American Philosophical Quarterly*, 29.
- . 1998. *The Authority of Reason*. Cambridge: Cambridge University Press.
- Larmore, Charles. 2008. *The Autonomy of Morality*. Cambridge: Cambridge University Press.
- Lewis, Paul. 2009. "Commitment, Identity and Collective Intentionality: The Basis of Philanthropy." *Conversations on Philanthropy VI*: 47-64.
©2009 DonorsTrust.
- Sen, Amartya. 1997. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Journal of Philosophy and Public Affairs*, 6.

COMMENT

WHAT IS WRONG WITH A SOCIAL PREFERENCE, OR TWO?

Shaun P. Hargreaves Heap

People act philanthropically, Paul Lewis argues, because these actions serve to fix or express their identity. To make sense of this, he sketches a model of individual agency that has a two-tiered structure in which what is added to a conventional instrumental model of acting to achieve an end is, variously, a sense of commitment or an ability to switch into a mode of “we reasoning.” In both cases it is membership in a group that creates the possibility of this distinct second motivational tier. He also argues that Hayek’s defense of the classical liberal view of social/market order depends on such a model of agency (and the implicit possibility of philanthropic behaviors).

I think Lewis is right on the matter of what model of agency can make sense of both philanthropy and Hayek. I shall engage with his argument, nevertheless, in two ways: first through a more detailed comparison with the standard “economic” or instrumental explanation of philanthropy and then with some related thoughts on the moral underpinnings of the market order.

A Comparison with the Standard Instrumental Explanation

The standard explanation of philanthropy in economics grants that individuals can have so-called “social” or “other-regarding” preferences. As a result, when they act to satisfy best their preferences (in the straightforward, instrumental manner) they can exhibit forms of pro-social or unselfish behavior. In this sense philanthropy is unmysterious, and the question arises as to why some more elaborate model of agency, such as that proposed by Lewis, should be required. The economic model would seem to have much to commend it, since much has been discovered recently about pro-social behavior in laboratory settings through the use of this model.

For example, we know from these experiments that (1) pro-social behaviors are really rather ubiquitous (that is, people act so in a variety of decision settings), but (2) these behaviors are distributed differently across the population (perhaps as much as half of the experimental population don't ever exhibit these behaviors), (3) they often have a reciprocal character and are sensitive to group allegiance, (4) they can be triggered in some institutional contexts but not others, (5) communication helps, and (6) considerations of fairness often dominate mere additions to individual welfare.

These are all rather useful regularities to know about pro-social behavior, and in large measure they can be inferred from behavior in the laboratory precisely because these other-regarding considerations are incorporated into the formal instrumental model of decision making. Without this encompassing formal model, the interpretation of behavior in these terms would be very much more difficult.

So, with this range of insight into the regularities of philanthropy, what is gained by having a more elaborate model of individual agency?

A first thought here is that the Lewis account has the advantage of making membership in a group quite crucial (indeed necessary) to the existence of pro-social behaviors, whereas it emerges as a contingent feature of these behaviors in the standard instrumental model. The necessity arises for epistemic reasons. People do like to reflect on what they do, and these reflections are a source of feelings of self-worth (or its reverse), but the standard for these judgments has to be external to the individual; otherwise, the judgments will be tainted by the suspicion that they are self-serving and so lose their psychological charge. Your group supplies this external standard. This may not at first seem terribly important, but it is. Groups have a discursive life. That is, members of groups talk to each other over what are good, bad, honest, honorable, just, etc., behaviors, and in this way the norms of a group which influence pro-social behavior adjust and evolve. But why might this be important?

There are two ways. First, when evaluating outcomes and forming policy prescriptions, one has to be concerned not just with how well any arrangement (or policy) satisfies people's preferences (as in the conventional welfare economics that is based on the instrumental model), one also has to attend to how the capacity of individuals to engage in the normative discussions which underpin these reflections of self-worth is served.

Second, it is important for explanatory reasons. The difficulty with the standard instrumental model from this standpoint has always been that it supplies no account of the origin of people's preferences. This is not a terrible weakness, at least in an explanatory sense, as long as people's preferences seem reasonably stable. Unfortunately, what the experimental evidence demonstrates is that these "social" preferences "come and go" in intensity depending on institutional context. In other words, they are sometimes endogenous rather than exogenous. For this reason we need some account of social preference formation, and this is where the Lewis-type model has an advantage, for the reason that I have just sketched.

There is, however, something more that we know about the endogeneity of preferences, which again can be related to Lewis-like models. It comes from the psychological theory of extrinsic and intrinsic reason and the earlier work in cognitive dissonance theory. In these theories, intrinsic reasons for action (which often correspond to the motivation that comes from a social preference, to use the language of the standard instrumental model) get crowded in or crowded out in reasonably predictable ways. The broad idea is that people like to find a reason for their action in an economical sort of way, and they have two generic types of reason that can be drawn upon for this purpose. There is extrinsic reason, which comes from the externally given circumstances of the action: one might buy something, for example, because the price is low. The price being right, so to speak, provides an extrinsic reason for the action. Alternatively, there is intrinsic reason: that is, an action is not contingently the best thing to do, but it is intrinsically the right thing to do. The "economical" part of this theory arises because people like to have a reason for action but only a sufficiency and not a surfeit. Thus if an action has both extrinsic and intrinsic reasons, it is overegged, and people adjust the reason they have control over, the intrinsic variety, downwards—there is a motivational crowding-out. Alternatively, if a person does something and there are weak or no apparent reasons, then the person adjusts the intrinsic reason for the action upwards—there is crowding-in.

I say this can be related to the Lewis-type model of agency because both his and the psychological literature share a view that people reflect on what they do. They like to make sense of what they do. In addition, the source of the problem when there is a surfeit of reason essentially turns on an epistemic issue that is

related to the one I sketched above. When an action is dually motivated as outlined in this psychological literature, there can be no unambiguous identification of the action with the relevant norm of intrinsic reason, and unless a norm can be instantiated reasonably unambiguously in action, it will wither.

There is another and increasingly popular explanation of the origins of social preferences which sits less comfortably with the Lewis-type model, which I should mention. It comes from evolutionary psychology. The idea here is that these social preferences are a genetic (or possibly also a cultural) inheritance of a set of behaviors that proved successful, in an evolutionary sense, in settings commonly encountered in the hunter-gatherer societies that have characterized most of human history. Thus we have some predispositions to act in these pro-social ways because there are some “modern” decision problems that are very close or similar to problems encountered in the *longue duree* of human history where such pro-social behavior proved evolutionarily valuable. In this way, philanthropy is not really thought about seriously at all; it is more like a tic, an involuntary, evolutionary twitch.

This is tempting for those wedded to the instrumental model, because it seems to explain the existence of social preferences and why they might appear more strongly in some circumstances than others without departing from that model in any deep sense. It is also capable of suggestive insights. For example, it has been argued that altruism within a group is most likely to arise when there is competition between groups and as a result altruism within a group may be associated with intergroup behaviors marked by hostility. Whether it can explain the variety of human behaviors in the experimental population in the *same* setting is less clear. What is clear, however, is that evolutionary accounts of *behavior* are just that. They are not an account of how we come to attach symbolic significance to these behaviors (describe them as “just,” “right,” “honorable,” etc.), and this is the fact of social existence that the Lewis-type model addresses. We are concerned with the meaning of our actions; we reflect on what we do, and we share in those reflections with others.

Market Society

The observation Lewis makes regarding how quite basic market transactions often rely on a background set of shared beliefs that people’s behavior will, in some relevant respects, be normatively constrained has often been a source of

concern. The worry arises because it seems that the expansion of the market might progressively undermine its moral underpinnings in two ways. First, from a sociological perspective it seems that moral norms might be generated from the socializing experiences of the retreating nonmarket institutions (and so might gradually disappear). Second, from psychology, because the market tends to be associated with the operation of extrinsic reason, it seems markets might crowd out intrinsic reason in the manner I have just sketched. I want to dispute the logic that turns these two tendencies into some internal contradiction of a market society.

In particular, I want to argue that what is likely to happen in market societies is not so much an evacuation as a change in the normative constraints on behavior. The reason for believing this comes from considering whether it is possible to imagine a society of people who are solely preference satisfiers capable of deriving a sense of self-worth from the activity of preference satisfaction. One might suppose that this would be possible if people had a higher-order preference for being preference satisfiers. This would appear to retain the simple model of preference satisfaction while gesturing in the direction of some kind of motivational reflection. The difficulty with this line of argument is that even if preference satisfiers have a preference for being preference satisfiers, this would not give them a reason or justification for valuing being a preference satisfier, because preferences are just that; they are not reasons.

A preference satisfier who derives a sense of self-worth from the activity of preference satisfaction has to have some expanded motivational wellspring that deals in reasons for behaving as a preference satisfier. Or to put this the other way around, if we observe people operating in the paradigmatically market mode of preference satisfaction and they also appear to be moved by considerations of self-worth, then they cannot *solely* be preference satisfiers. They must be something more in the sense that they are plugged into a normative discourse that valorizes preference-satisfying behavior. One has only to consider one obvious candidate element in such a discourse—the belief in freedom, or liberty—to see that this would still involve some constraints on individual behavior because individual liberty is never absolute: it is always practically and politically defined as the freedom to do anything that is consistent with others in society having the same freedoms.

It is interesting that experiments are beginning to throw some light on how markets affect behavior along these lines by triggering different normative

constraints. I say this at the end by way of reinforcing my first observation: the standard model has much to recommend it even if it cannot be the whole story. But as Paul Lewis says at the end of his paper, and by way of endorsing what he says, the story of the market is one for another time.