

3

Studying Scientific Virtues

Bridging Philosophy and Social Science*

Robert T. Pennock and Jon D. Miller

Studying the Scientific Character Virtues

Philosophers of science have tended to focus on philosophical issues involving scientific concepts in particular fields, such as physics or biology, and scientific reasoning in general, such as the nature of scientific explanation and the justification of inductive inference. Comparatively little attention has been paid to the character of the scientist and the identification and investigation of traits that make for excellence in scientific practice, especially when these are conceptualized as having a moral structure. One strand of Pennock's philosophy of science research has focused on scientific methodology, including confirmation, methodological naturalism, and model-based reasoning; a second strand of empirical research considers the evolution of intelligent behavior. These two strands have been linked by an interest in the relationship of epistemic and ethical values in science, which led him to turn to questions about character and the scientific mindset and to develop an account of scientific excellence and its development that combines Aristotelian virtue ethics and Darwinian evolution (e.g., Pennock 2002, 2015, 2019).

When virtue ethics has been applied to science previously, it has mostly been in terms of how general human virtues may be relevant in science. Pennock's alternative approach, which starts with a narrower notion of vocational virtues, however, focuses on character traits as they relate directly to the central, guiding purpose of different vocations. For the practice of science, this approach suggests that there are particular virtues, from veracity and curiosity to objectivity and skepticism, that should be given special weight for fostering excellence in scientific practice and scientific flourishing generally. Over the last couple of decades, Pennock has developed this view philosophically and deployed it for various applications (e.g., Pennock 1996, 2001, 2018; Pennock and O'Rourke 2017). For the most part, it would have been possible to pursue this line of research entirely with a philosophical methodology, but adding a social science component was always an enticing possibility.

From the point of view of sociology of science, an empirical study would provide a look into the culture of science that complemented and extended Robert Merton's work, which had focused on institutional values, setting aside more specific questions about ethical perception and motivation of scientists themselves. And from the point of view of philosophy of science and ethics, such an interdisciplinary study of the ethical perceptions of exemplary scientists held the promise of providing a kind of test of some aspects of this virtue philosophy of science. This kind of investigation required collaboration with a social scientist, which led to the research partnership with Miller and a study of scientists' views about the virtues of their vocation that is now in its sixth year. In this chapter, we will reflect upon some of what we have learned in attempting to bridge these two disciplines and the unique value that doing so has for understanding a cross-cutting topic such as scientific virtue. Before describing our own work, however, we present a couple of cases to show why collaboration between philosophy and social science is infrequent and what might account for the general separation between these disciplines.

When Philosophers and Social Scientists Collaborate

Psychology and the other social sciences began as branches of philosophy, but since the mid-nineteenth century they have gone their own ways, and practitioners in these fields for the most part now operate independently of one another. When philosophers and social scientists do meet, as occurs periodically for scholars in neighboring areas of research, relations are cordial, though their differences can sometimes make it difficult for them to work together on interdisciplinary projects in substantive ways. It is common to ascribe such difficulties to differences in their language. Indeed, in their conversations at interdisciplinary research conferences that they might both attend it does seem at times that scholars are talking past each other. Sometimes they use the same terms but in different ways, which can make a conversation especially confusing. However, their differences are not solely a matter of language; their methods and practices are mostly distinct as well.

are mostly distinct as well.

Consider the role of data in their respective research about ethical norms. It is not just that ethicists rarely speak of data: they often appeal to ethical intuitions in their arguments, but they do not usually think of these as data and only occasionally use data in a way that social scientists would recognize. If data do appear as part of an ethical argument, they are mostly just part of some background secondary premise and so are taken for granted, as in a utilitarian argument. Why? Because philosophers are interested in figuring out what is proper or appropriate, and this kind of normative question is investigated

more commonly by means of analysis and conceptual argument than by empirical observations.

Similarly, it is not just that social scientists rarely talk about values. They may investigate social norms, but not in the evaluative way that philosophers deal with what they call norms. Why? Because social scientists are more interested in describing values than evaluating them. Most social scientists insist upon being scrupulously neutral with regard to different values they may be describing—for example, values from different cultures. It is not their place or interest to make judgments about the values of the people they study.

That ethicists and social scientists' interests are mostly different means that they have little reason to pay attention to what is happening in the other's domain. Their professional literature rarely has the need to cite work from the other, so philosophers are mostly unfamiliar with the social science journals and social scientists rarely dip into the philosophical literature. With their eyes on their own acreage, workers on both sides of the disciplinary wall are thus often unaware of internal conflicts going on in the neighboring field. Ethicists usually have little reason to concern themselves, for instance, with differences between cognitive schema theorists and trait theorists or other developmental camps in psychology; if they notice them at all. Psychologists, similarly, for the most part have little knowledge of, or interest in, debates among ethical theorists about, say, the pros and cons of virtue ethics vis-à-vis deontology or consequentialism.

One finds a similar pattern as one moves from psychology to other social sciences. This is not to say that social scientists have been entirely uninterested in differences already noted, the most common sort of studies conducted have been ones that investigate scientific misconduct and attempt to ascertain how common various kinds of questionable research conduct are (e.g., John, Loewenstein, and Prelec 2012). This is still a descriptive question, though in this case a prior specification of what kind of behavior is unethical in science is required. For such a study to get off the ground, social scientists must work with ethicists, at least up front. This is perhaps the simplest form of collaboration—research ethicists say what behaviors are unethical and social scientists then can get on with their basic form of measuring the incidence of that behavior. For many purposes, this clearly beneficial to get real data about the prevalence of various forms of misconduct. For this kind of research, social scientists need not concern themselves about the ethical theory or reasoning behind the identification so long as they have some well-defined behavior to measure.

As in this example, relatively few social scientists need to or care to derive further into ethical thought itself. Robert Merton's pioneering work on the ethos of science is a notable exception, but even Merton paid little attention to

philosophical issues and mostly discussed the scientific ethos by simply identifying institutional norms and then analyzing their utility for maintaining social cohesion and behavioral compliance (Merton 1938, 1973). On occasion, however, there have been scholars who have embarked on an investigation on the border that took them beyond the current division and back to a place where the disciplines shared some common interests.

Lawrence Kohlberg, who brought investigation of morals back into psychology, is one notable example (Kohlberg 1969). Using a Piagetian model of developmental stages, Kohlberg studied how people reasoned about moral issues, and in his data he identified six stages (grouped into three levels) that individuals pass through in the development of their moral reasoning. In the earliest, "pre-conventional" level, children first reason in terms of obedience and punishment (thinking of how to avoid punishment), followed by a self-interest orientation (thinking in terms of "what is in it for me" and how to receive rewards). Beginning in adolescence, a second level of "conventional" moral reasoning develops, first with an emphasis on trying to conform to social expectations, and later with a "law and order" mentality in terms of societal authority and social conventions. In the "post-conventional" level, more abstract reasoning develops, determined by some joint process for decision for the common good, and then, in the sixth stage, in terms of universal ethical principles, especially of rights and justice.

Kohlberg's moral stage theory in the psychology of moral development has been influential, but its presentation emphasized certain ethical views over others in ways that some saw as problematic. The highest, sixth stage seemed to place deontological ethics (exemplified by Kant's view of morals as universal laws) over utilitarian ethics, which were placed earlier, in the fifth stage. Carol Gilligan's *In a Different Voice* (1982) suggested that it overlooked a distinctively female ethics of care. (We will look at that criticism in a bit more detail shortly.) There was certainly a blind spot with respect to virtue ethics, as pointed out by David Carr, who thought that was all the worse for the whole Kohlbergian approach, especially with regard to implications for moral education (Carr 1996). Kohlberg later adjusted and expanded his model in response to some of these points. When one is researching a topic such as moral psychology, it becomes necessary to talk with one's neighbors.

In a paper about their own work on "psychologized morality," partially inspired by Kohlberg's research, the psychologists Daniel Lapsley and Darcia Narvaez reflect on some of the issues that arise for workers at the border between psychology and ethics (Lapsley and Narvaez 2008). We would like to now add to their useful reflections for a slightly broader notion of social science, and we will adopt their use of the metaphor of collaboratively repairing the wall between

these fields from Robert Frost's poem "Mending Wall" (Frost 1914) as a way to frame the issue of the nature of interdisciplinary collaborations of this sort. There is much that could be said about the relationship between these two neighbors, but we want to make just a few points about the relationship between data and values in such collaborations.

Theory: At the Descriptive/Prescriptive Divide

A standard theoretical reason for the separation between ethics and science involves the problematic conceptual gap that philosopher David Hume identified between facts and values: that something is the case does not tell us that it ought to be (Hume 1739). The philosopher G. E. Moore called one form of this leap from the former to the latter (in particular the reduction of a moral term such as "the good" to some natural property such as the pleasurable) the "naturalistic fallacy" (Moore 1903). It is easy to see how this problem applies to the sort of interdisciplinary research we are considering. Social science investigates factual questions and provides descriptive information. Ethical philosophy investigates value questions and provides prescriptive evaluations. Both sorts of inquiries are sometimes referred to as "normative," which adds confusion to what is already a notoriously complex issue, but one basic way to put the point is as follows: No matter what a descriptive study finds about community C's or subject S's ethical beliefs, ethics may still properly ask whether C or S should or should not hold those beliefs. Thinking that mere facts about a culture's or subject's views are sufficient to determine ethics leads to the conceptual incoherence of cultural relativism or subjectivism. This is Philosophy 101. Kohlberg took this issue seriously and wrote a major article about how his empirical work could avoid the pitfalls of the naturalistic fallacy (Kohlberg 1971).

This distinction is the most plausible reason why philosophers and social scientists are happy to stay in their respective fields with a wall that marks a dividing line between their areas of interest. But what is the nature of this wall? If it is taken to be an impermeable barrier, then one might wonder what use information gathered from science generally and social science in particular could ever have for ethics. On the other hand, if it is taken to mark a strict separation, then it might serve as a regulative principle for both philosophers and social scientists, making it clear when one has improperly trespassed on the other's property. We may gain some insight into the nature of the wall by looking at a couple of cases of border crossings that were occasion for both friendly and more strained interactions between the neighbors.

An especially interesting border case involved issues raised by Kohlberg's research. His own work on stages of moral development became a topic of

discussion in philosophy courses, but it did not have nearly the impact as the work of Carol Gilligan, who began as his research assistant but whose own research led her to identify what she purported was a distinctively different, "female" moral voice that reasoned in terms of care for others and preservation of relationships compared to a "male" focus on rights and fairness. Here is a case where empirical data from social science seemed to directly influence ethicists, where empirical data from social science seemed to directly influence ethicists. Gilligan's work led to suggestions for reappraisal of standards of ethical judgment, first among feminist philosophers and then more broadly. Largely because of the influence of her empirical work, it became and remains common to include discussion of care-based ethics in ethics textbooks.

Not all philosophers embraced Gilligan's work wholeheartedly, however. Ethicist Owen Flanagan warned of philosophical dangers he saw in her work and Kohlberg's as well, especially with regard to claims about particular hierarchies of moral adequacy, which presumed resolution to disputes among different ethical theories that had by no means been accomplished. Kohlberg's account drew heavily from the influential work of ethicist John Rawls, but important as that theory of justice was and remains, it seemed premature to crown its notion of morality as more justified than utilitarian or other ethical accounts. Divisive arguments about whether "male" justice or "female" care is superior bear witness to how psychological claims can negatively affect moral discourse. Flanagan recommended that discussion of possible differences between how women and men view morality ought better be pursued "in a scientific context freed of the ideological weight of moral stage theories with their tenuous suite of claims about universality, irreversibility, and, especially, adequacy" (Flanagan 1982, 512).

Philosopher Debra Nalis was harsher, arguing on both philosophical and empirical grounds that Gilligan's work was "fallacious and unreliable" (Nalis 1983). She was especially critical of reifying psychological scales of "moral maturity," but she also delved into various methodological problems with the interview process, calling into question Gilligan's conclusions about differences between male and female moral standpoints. Subsequent empirical work bore out this criticism. Other researchers were not able to replicate Gilligan's results, and a meta-analysis of research on gender differences in moral orientation showed at best a weak relationship between care and justice orientations in women and men (Jaffee and Hyde 2000). Nalis did point out that Gilligan had noted at the very beginning of *In a Different Voice* that the different voice she described was characterized by theme, not gender, but that important caveat was mostly lost in the book itself and in the subsequent reception of the work. Nalis concluded with a warning: "This type of research is social science at sea without anchor, and no one is out of danger" (Nalis 1983, 664).

Now we turn to a second example—one where the challenge comes from the other direction. This case involves what is known as the "situationist critique" of

virtue ethics. We have already mentioned the debate between virtue ethics and theories of deontology and utilitarianism as an example of an internal topic of interest among ethicists. Virtue ethics, which derives from Aristotle, especially his *Nicomachean Ethics*, and has seen a resurgence of interest in recent decades, focuses on the notion of moral character. Rather than dealing primarily with the morality of actions and events as deontological and utilitarian ethics do, virtue ethics deals first with agents and states of agents. In an Aristotelian framework, virtue ethicists analyze how the development of moral character leads to human flourishing. If a person is courageous, for instance, that trait may be expected to motivate appropriately courageous behavior when circumstances warrant it. Similarly, having the virtue of generosity ought to lead one to act generously in appropriate situations. Aristotle described various ways that virtues can be developed, such as by learning from other individuals who have acquired practical wisdom.

Deliberation about the justification of ethical theory mostly involves conceptual arguments, but in an unusual case this standard account of virtue ethics was criticized on empirical grounds using data from psychology. The philosopher Gill Harman, pointing to empirical studies that highlighted the problem of explaining behavior in terms of distinctive characteristics of an agent over relevant details of the agent's perceived situation, called the virtue ethicists' appeal to moral character traits an example of a "fundamental attribution error" (Harman 1998–99, 329–30). Ethicist John Doris extended this critique, highlighting an experiment from social psychologists Alice Isen and Paula Levin (1972) that showed how helping behavior varied as a function of mood state—that is, of "feeling good" in a situation. In their experiments, something as simple as finding a dime in the coin return of a public phone led to a statistically significant increase in helping behavior. Doris cited this experiment as representative of the "situationist" tradition in social and personality psychology, which rejects explanations of behavior in terms of standard theoretical constructions of personality, and he concluded, like Harman, that it calls into question neo-Aristotelian virtue ethics based on the idea of steady character traits (Doris 1998, 504). Such experiments suggest, he argued, that one cannot reliably predict helping behavior based on traits such as selfishness or compassion (as in his hypothetical example of subjects in the *dime study*) and that an alternative conception of moral personality and of ethics is required that does not posit reliable moral dispositions.

The descriptive/normative divide notwithstanding, this situationist challenge was recognized by ethicists as a real issue for neo-Aristotelian virtue ethics. One may of course say, as Doris himself acknowledged, that showing the descriptive inadequacy of an ethical theory is not sufficient for showing that it is normatively inadequate, but the descriptive data do bear upon its plausibility, which led to a lively debate in the literature. We mention just a couple of examples.

Diana Fleming confronted the situationist challenge from both sides, questioning the empirical evidence and its relevance. Experiments such as Isen and Levin's, she pointed out, are "poorly suited to detecting the existence or influence of traits" (Fleming 2006, 38), as they focus on particular situations rather than behavioral consistency over long periods of time and thus "bias the evidence in favor of situational factors" (38–39). She also rightly pointed out that it was a fundamental mistake to think that character traits should operate "without regard to environmental factors," and she argued that once this is acknowledged, it is easy to see that virtue ethics has the conceptual resources to build in "situation sensitive" traits (41).

Nancy Snow responded to the situationist challenge in even greater detail in a book-length account that defended an account of virtue as "social intelligence" (Snow 2010), with extensive discussion of social science research to ground her theory. Contra arguments that situationist experiments undermine the notion of stable character traits, she pointed to other social science research showing how agents' responses to situations also depend upon how they construe the situations they find themselves in. This adds another dimension to moral behavior; it depends not just on dispositions and environments but also on construal. In addition to mood-effect studies such as those of Isen and Levin, Snow addressed a range of others including the infamous Stanford Prison Experiment and the Milgram obedience-to-authority experiments, concluding that situations have meaning for people and that subjects' moral behaviors must be judged with that factor in mind (Snow 2010, ch. 5).

It is not our intention to get into the details or take a stand on the specifics of these debates. We outline them here for the way they reveal the various sorts of conversations that may arise at the border between philosophy and social science. They also illustrate occasional conflict between these parties, though that need not be inevitable; confusions and potential conflicts may be sorted out if each side brings its own expertise to the conversation and does not presume to discount that of the other.

Kohlberg may have neglected virtue ethics, but we are not as pessimistic as Carr in thinking that any such view is incompatible with virtue ethics. The idea of a universal morality surely does not preclude the place of moral judgment values the goal of learning from the practically wise how to balance complex values when considering what is ethical in particular cases. Moreover, even though the empirical basis of Gilligan's claim of a distinctively female ethical voice seems to have been mostly undermined, her idea of an ethics of care remains of philosophical interest. Future research may shed light on whether there is any psychological import to the small empirical differences between men and women with regard to care-based versus justice-based moral reasoning, but no matter how this turns out, ethicists could still argue about how to balance justice and care in moral reasoning and decision-making.

Situationists may be wrong in their specific challenge, or at least premature in their conclusions, but they could have been right, and there are other empirical data that may be relevant. For instance, if moral psychology were to show that human beings are psychologically incapable of being virtuous or behaving virtuously, then that would seriously reduce the viability of applying virtue ethics to human morality. Metaethics correctly notes that ought implies can, which means that we at least need to be capable of virtue if it is to be recommended as the correct moral account for human beings. It is extremely unlikely that we are incapable of virtue, but the same point holds for weaker possibilities. The general point for us here is that the entire debate occurs because everyone agreed that for some important aspects of virtue ethics, the empirical data from social science is relevant. As Snow rightly noted, "Philosophers have much to learn from studying the work of psychologists" (Snow 2010, 117).

What these cases show is that there are interesting and substantive points of contact between descriptive psychology and prescriptive virtue ethics where each may reasonably influence the other. These and similar debates show that social science and ethics are not entirely independent and that interesting things may happen where the fields meet. There is much to talk about over the wall, including questions about the wall's proper placement.

When conflict arises between philosophical and social science perspectives, it could easily turn into a war or at least a standoff if one thought that the issue involves which side is more fundamental and gets to decide the answer. But this is usually the wrong way to think about such cases, as the sides are mostly interested in different questions, as we saw, and conflicts are mostly the result of miscommunication about their aims. Clarifying not just terms but methodological assumptions can go a long way toward maintaining peaceful relationships at the border. As we have also seen, however, in other cases there may be real points of contact where some questions do touch substantively on both sides. In such instances, neighborly relations require a more cooperative approach.

While it may offer data about moral development, social science is worth taking seriously only if it has constructed its instruments and protocols so that we can trust that what they are measuring is ethically significant, rather than some related but non-moral trait. Are moral psychologists (and other social scientists) really telling us something about virtue? Any study needs to begin with the appropriate theory (here an ethical theory) before it can state meaningful hypotheses to test. This requires the expertise of philosophers. For instance, an ethicist would need to first explain what being courageous means in virtue-ethical terms, say, as opposed to what it means colloquially, before psychological studies can investigate how to cultivate it as a moral trait. A study that failed to appreciate the difference between the philosophical notion of a character trait and the psychological notion of a personality trait is likely to confuse rather than enlighten.

In the other direction, when studies are appropriately relevant, ethicists must be constrained by what social science finds out about limits of human ethical behavior, or other facts that bear upon one or another aspect of ethical thought and behavior. In such cases, data may make a real difference and could even play a role in assessing different ethical theories. For example, virtue ethics claims an advantage over other ethical theories in that moral motivation is supposedly built into its view. It could turn out, however, that social science finds that ethical motivation works differently than theory predicts—perhaps, for instance, training in a rights-based theory actually is better than training in a virtue-based approach at leading to moral action. Again, this is a hypothetical scenario, but if empirical data substantiated that possibility, then it would be incumbent upon ethicists to consider its implications and rethink their views.

These are just a few examples where collaboration between ethicists and social scientists could prove to be fruitful for both. We now return to our own collaboration, which we believe is another case where an interdisciplinary approach opens up new ways of thinking about some old questions.

Our Scientific Virtues Survey Collaboration

Pennock's philosophical theory of vocational virtue holds that practices and methods that characterize a discipline are justified in relation to its central guiding purpose and general human values, and that the character virtues involved in the excellence and integrity of that vocation are those traits that ought to cultivate for it to flourish. With regard to science, which aims to discover truths about the natural world, the theory thus holds that scientists ought to value traits such as veracity, curiosity, objectivity, and related virtues; these would be expected to constitute the scientific mindset, helping science flourish. To see what ideals scientists themselves actually held, in the late 1990s Pennock created a preliminary questionnaire and began a series of informal interviews with scientists. This provided anecdotal support for the account, but it wasn't until the early 2010s that the opportunity arose to develop and mount a systematic, formal, national study in collaboration with Miller.

Miller's work has focused on empirical descriptions of adolescent and adult development of life plans and goals, including occupational objectives and the acquisition of skills and information needed to achieve various life goals (e.g., Miller 1983, 2000, 2010). This life course approach is reflected in two major strands of work.

In 1987, Miller launched the Longitudinal Study of American Youth (LSAY) to study the development of student engagement with science and mathematics and the impact of these early learning experiences on the selection of career

paths and life goals.¹ The study adopted Berger and Luckman's developmental perspective (Berger and Luckman 1966), which minimizes genetic explanations and argues that newborn children are blank slates that are shaped by their family and life course experiences. Language plays a central role in the definition of reality and the transmission of values. In this model, a child born in Ann Arbor, Michigan, and transported to Beijing within a few days of birth and raised by a Chinese family would be expected to grow up speaking Mandarin and to adopt the values and norms of his or her significant others (the adults who nurture and raise the child). Conversely, a child born in Beijing and transported to Ann Arbor within the first few days of life and raised by a midwestern American family would grow up speaking American English and would adopt the values and norms of his or her American significant others.

Applying this model, the LSAY collected questionnaires and achievement tests for its 5,900 participating students several times each year and extensive questionnaire information from each student's science and mathematics teachers. One parent of each child was interviewed by telephone each year during secondary school. Extensive school and community context information was added to each student's record. After secondary school, the LSAY followed each student to work or postsecondary study, using a combination of mailed questionnaires and telephone interviews. In 2007, the LSAY changed to a system of online data collection for all respondents who were willing to respond online and to mailed print questionnaires for other respondents. The LSAY continues to follow 5,100 of the original 5,900 respondents (the eligible sample who are still alive, living in the United States, and mentally and physically able to complete a questionnaire). The second strand of Miller's work is a forty-year time series of adult studies that measure patterns of adult issue interest, scientific literacy, and democratic engagement in the formulation of science and technology policy in the United States. The insights gained from the developmental analysis in the LSAY allow the construction of questions for cross-sectional respondents that provide useful development indicators (obtained by recall). But the selection of a new probability sample of American adults each year allows the construction of time trends over decades.

At one level, all of this work is empirical and descriptive, but it also reflects a commitment to democracy and democratic values. A good deal of Miller's work has been based on his measurement of civic scientific literacy—defined as the level of scientific construct understanding necessary for a citizen to engage in science and technology policy discussions. Democracy depends on meaningful citizen participation, which implies a level of issue understanding sufficient to make policy-relevant judgments in voting and in contracting or pressuring electing representatives. A second set of value issues pertain to educational and social disparities in society. One of the results of consistent measurement over

a period of decades is that it is possible to see patterns of differentiation among individuals and groups. Miller's work has focused attention on the growing disparity in educational opportunity and attainment in the United States and its consequences for employment, health, and quality of life. Here the distinction between description and prescription does apply, as the data do not by themselves prescribe a specific solution or even a precise definition of a problem. However, they do provide a common ground for discussion.

Both strands of this social science research were relevant to collaborating on an interdisciplinary investigation of scientific character virtues and values. We began our study in 2012 with pilot interviews to develop our survey instrument and kicked off the full national study in 2014. Drawing a random sample from the population of peer-identified exemplary scientists (e.g., elected fellows of the population of peer-identified exemplary scientists (e.g., elected fellows of the major scientific societies) and a population of early-career scientists, we then solicited and conducted phone interviews, which averaged forty-five minutes. Subjects who did not want to do an interview could provide written answers. By 2018 we had data from more than 1,100 scientists.

There are many things that one could say about the nature of this collaboration, but this is not the place to discuss the details of the protocol or the results. Here we just want to highlight three points about our study that are illustrative of some of the methodological and other issues that we think are of interest for philosophers and social scientists who walk the wall at the border between their fields.

First, the core of our survey is not about what virtues scientists have but what ones they most value. This helped us get at the questions that are of philosophical interest. We did not ask about what character virtues/trait scientists do or do not exhibit; rather, we asked what traits were important (i.e., valuable) for science. This is asking for their normative assessment (i.e., their evaluative, ethical judgment) of possible traits.

In this way, our survey has a prescriptive focus on what scientists think they should aspire to, rather than a descriptive focus on what traits they may or may not display. It would be a very different kind of study to investigate how scientists are actually behaving. That kind of descriptive information is useful for other purposes, of course. Other studies have tried to investigate how widespread scientific misconduct (for instance, fabricating data) is, but for us the question of how many scientists depart from the virtue of honesty is secondary to the question of how important they think that value is to science in the first place. For many social scientists, this may not be a distinction that makes much difference for their purposes—it is just one measure of attitude among any number of others. But for philosophers, ideals and aspirations are normative in more than a descriptive sense, as they take us into the realm of the asking and giving of reasons. This is grist for the philosophical mill.

There are other aspects of our study, however, that are familiar to social scientists but will seem foreign to most ethicists. We turn now to two related examples.

For one set of questions in our instrument we asked subjects to give a quantitative rating of a list of virtues. For each virtue, we asked them to assess its importance in science by rating how essential the trait was to being an exemplary scientist on a scale of 0 to 10, with 0 being not at all essential and 10 meaning that it is very essential. This is not the sort of thing that ethicists would typically care or think to ask. Quantitative assessments do come into play in utilitarian arguments, where doing the right thing (i.e., that which produces the greatest good for the greatest number of people) requires that we assess the effects on happiness (or pleasure) that different choices would have. In reality, however, it is more likely to be economists who actually collect and analyze such data; utilitarian ethicists work out the theory but do not usually themselves apply it in this way. So, besides that, what possible philosophical value is there in asking for a quantitative assessment of the importance of some value/virtue?

One way to begin to answer this is to consider the alternative. For instance, take any random virtue and one could likely find a way that it could be useful in some particular scientific investigation. Take generosity; for example. An advocate for this virtue could easily find or concoct a case where it would behoove scientists to be generous—in the sharing of samples, for instance. If the only question were whether a virtue could be relevant for a scientist, one could come up with possible scenarios without troubling to run a survey. But this is where data can be revealing: it would be interesting to find that generosity (or some other common virtue) does not register as very important for scientists (or that some uncommon virtue does). We may certainly admit that generosity is a general virtue for human society without thinking that it should be applied with equal force in all human vocations compared to other virtues. This is why moral judgment is important after all.

Quantitative information from exemplary scientists (or some other vocational practitioners) provides a broad measure of the judgments of practiced researchers—those who presumably have acquired some practical wisdom from their experience. This kind of data is surely relevant to the assessment of the claims of virtue ethics, and it is better than a simple appeal to the intuitions of one or a few ethicists.

Moreover, such quantitative data allows one to ascertain whether there are patterns of emphasis. It could turn out, for example, that exemplary practitioners of other vocations all agree that everything on our list, or some other, were indeed virtues (i.e., that it is generally a good thing to be honest, curious, objective, and so on, or generous, caring, and chaste), which is by no means certain, but that would not be the end of the matter. If one asked only whether something was a virtue or not, one might have to leave it at that, but quantitative ratings allow

one to see whether there are patterns of weights that practitioners of different vocations give to different virtues. Pennock's vocational virtue account predicts that there ought to be such differences, depending in large part on what is the central, guiding purpose of the practice. If exemplary practitioners of different vocations all gave the same weights to the same set of virtues, then that would be a good indication that the account is wrong.

Of course, we did not stop with numerical ratings; we also collected a wide range of qualitative information. We asked scientists to explain the reasoning for their highest scores. We asked them to describe examples of how these appeared in practice. We asked them how they learned about these values and how they tried to pass them on to their own students. We asked them about scientific vices. We asked for anecdotes and stories that illustrated the nuances of these views and the reasoning behind them. Philosophers rarely collect such information in a systematic fashion: their method relies more on conceptual analysis, where a single case—sometimes even a hypothetical case—can provide fodder for extended reflection, discussion, and debate. Imagined examples can be as powerful as real ones when one is trying to work out how to think about what we ought or ought not to do in difficult ethical dilemmas, so it may not matter whether the examples were drawn from life or not. Because philosophical thought experiments are the bread and butter of much philosophical deliberation, many philosophers will have little use for even this sort of qualitative empirical data. However, it is the quantitative data and how to incorporate it into philosophical analysis that will require more significant rethinking of their usual modes of reasoning. This brings us to a third point of methodological interest that involves a specific, technical method of quantitative reasoning that we used.

Confirmatory factor analysis (CFA) is common in the social sciences, but it is probably fair to say that our study is the first time that it has been part of a philosophical investigation. The collection of the relative importance of each of several virtues provides a useful empirical basis for analyzing whether they cluster or are likely to be associated. A CFA procedure uses the variance or covariance among a set of variables (our measures of the importance of each virtue) and asks whether all of the measured virtues are of equal importance or whether some are distinctively more important and more likely to be selected together by exemplary scientists. In drawing descriptive conclusions, social scientists pay attention to critical methodological principles from the philosophy of science about the importance of subjecting hypotheses to rigorous tests, and CFA is one way to do this for certain kinds of models.

We see this examination of whether items (virtues) cluster or covary as a test of one aspect of the hypothesized view of scientists' values. If we start with the premise that all of the virtues included in our final questionnaire (which were

selected through pilot testing of a much larger number of prospective virtues) are equally likely to occur in a final cluster of virtues, then we would expect a uniformly strong (or high) set of factor loadings. But we found that only eight of the ten virtues we asked about in the final questionnaire formed a cohesive cluster and that two did not. We would characterize this finding as disconfirming of the central role of the two virtues that did not load on the common factor.

This finding does not mean that the two virtues that did not load on the common factor for our sample of exemplary scientists would not load on a similar factor for another sample of exemplary scientists or for similar samples of engineers, bankers, or other professionals. Neither does it mean that these virtues may not be important in some other way or for a subset of research. For instance, we hypothesize that the virtue of being collaborative did not load together with the others because subjects thought of significant cases of exemplary science that were the work of a scientist mostly working independently.

If hypotheses or expectations are disconfirmed in a number of examinations of this kind, scientists and social scientists would conclude that the virtues in question do not belong as a part of a core set of virtues. Ethicists should take heed as well, for this may be revealing of some unrecognized feature of the moral landscape of science.

Conclusion

In Frost's poem, two neighbors meet in the spring to repair their common wall. One notes that a wall is not really needed, as "he is all pine and I am all apple orchard," but the other says that good fences make good neighbors, so together they close any gaps that have formed over the winter. In their reflection on the neighboring fields of ethics and developmental psychology, Lapsley and Narvaez suggested that a fence is good "not because it keeps neighbors apart and distant but because the occasion of repairing it gives them reason to work together in partnership" (Lapsley and Narvaez 2008, 280). In this chapter we have examined a few elements that are broader field of social science, focusing especially on the relevance of empirical data. As we have seen, philosophers and social scientists do have different primary interests and these differences of interest may sometimes lead to their talking past one another, as is often the case for many interdisciplinary conversations. However, many such problems can be avoided once the two groups recognize and articulate the relevant differences. Difference of interests should be acknowledged (that is what the mending of the wall involves) so that neither side inadvertently insists that only their interests are legitimate (it would be a mistake to root out the apple trees to plant more pines, or vice versa).

We have also identified a few ways in which each neighbor's interest may properly intersect with the other's: where data and values may meet; and where practitioners on both sides may find their perspectives broadened by interaction with the other. In such cases, philosophers and social scientists may collaborate for their mutual benefit. Putting the point poetically, we may conclude that in mending walls between fields of apples and pine, it is perhaps worth considering the value of leaving a gap, perhaps with a gate, that will allow a way for the neighbors to cross over and be wide enough so that "even two can pass abreast."

Notes

- * This material is based upon work supported by grants to Pennock and Miller by the John Templeton Foundation under Cooperative Agreement Nos. 36630 and 42023, and to Pennock by the National Science Foundation under Cooperative Agreement No. DBI-0939454. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the John Templeton Foundation or the National Science Foundation.
- 1. Originally funded by the NSF for twenty years, the study is now funded by the National Institute on Aging and has been renamed the Longitudinal Study of American Life.

References

- Berger, Peter L., and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Anchor Books.
- Carr, David. 1996. "After Kohlberg: Some Implications of an Ethics of Virtue for the Theory of Moral Education and Development." *Studies in Philosophy and Education Society of Moral Education and Development* 32, no. 4: 504–30.
- Doris, John M. 1998. "Persons, Situations, and Virtue Ethics." *Notas* 32, no. 4: 504–30.
- Flanagan Jr., Owen J. 1982. "Virtue, Sex, and Gender: Some Philosophical Reflections on the Moral Psychology Debate." *Ethics* 92, no. 3: 499–512.
- Fleming, Diana. 2006. "The Character of Virtue: Answering the Situationist Challenge to Virtue Ethics." *Ratio* 19, no. 1: 24–42.
- Frost, Robert. 1914. *North of Boston*. New York: Henry Holt and Company.
- Gilligan, Carol. 1982. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.
- Harman, Gilbert. 1998–99. "Moral Philosophy Meets Social Psychology." *Proceedings of the Aristotelian Society* 99, no. 1: 315–31.
- Hume, David. 1739. *A Treatise of Human Nature*. London: John Noon.
- Iken, Alice M., and Paula F. Levin. 1972. "Effect of Feeling Good on Helping: Cookies and Kindness." *Journal of Personality and Social Psychology* 21, no. 3: 384–88.
- Jaffee, Sara, and Janet Shibley Hyde. "Gender Differences in Moral Orientation: A Meta-Analysis." *Psychological Bulletin* 126, no. 5: 703–72.

- John, Leslie K., George Loewenstein, and Drazen Prelec. 2012. "Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling." *Psychological Science* 23, no. 5: 524-32.
- Kohlberg, Lawrence. 1969. "Stage and Sequence: The Cognitive Development Approach to Socialization." In *Handbook of Socialization Theory*, edited by D. A. Goslin, 347-480. Chicago: Rand McNally.
- Kohlberg, Lawrence. 1971. "From Is to Ought: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development." In *Cognitive Development and Psychology*, edited by T. Mischel, 151-235. New York: Academic Press.
- Lapsley, Daniel K., and Darcia Narvaez. 2008. "Psychologized Morality' and Ethical Theory, or, Do Good Fences Make Good Neighbors?" In *Getting Involved: Global Citizenship Development and Sources of Moral Values*, edited by Fritz K. Oser and Wil Veugelers, 279-91. Rotterdam: Sense.
- Merton, Robert K. 1938. "Science and the Social Order." *Philosophy of Science* 5: 321-37.
- Merton, Robert K. 1973. "The Normative Structure of Science." In *The Sociology of Science: Theoretical and Empirical Investigations*, 267-78. Chicago: University of Chicago Press.
- Miller, Jon D. 1983. "Scientific Literacy: A Conceptual and Empirical Review." *Daedalus* 112, no. 2: 29-48.
- Miller, Jon D. 2000. "The Development of Civic Scientific Literacy in the United States." In *Science, Technology, and Society: A Sourcebook on Research and Practice*, edited by David D. Kumar and Daryl E. Chablin, 21-47. New York: Plenum.
- Müller, Jon D. 2010. "The Conceptualization and Measurement of Civic Scientific Literacy for the 21st Century." In *Science and the Educated American: A Core Component of Liberal Education*, edited by John G. Hildebrand and Jerrold Meinwald, 241-55. Cambridge, MA: American Academy of Arts and Sciences.
- Moore, G. E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Nails, Debra. 1983. "Social-Scientific Sexism: Gilligan's Mismeasure of Man." *Social Research* 50, no. 3: 643-64.
- Pennock, Robert T. 1996. "Inappropriate Authorship in Collaborative Scientific Research." *Public Affairs Quarterly* 10, no. 4: 379-93.
- Pennock, Robert T. 2001. "The Virtuous Scientist Meets the Human Clone." In *New Ethical Challenges in Science and Technology: Sigma Xi Forum 2000 Proceedings*, 117-24. Durham, NC: Sigma Xi.
- Pennock, Robert T. 2002. "Research Funding and the Virtue of Scientific Objectivity." *Academic Integrity* 5, no. 2: 3-6.
- Pennock, Robert T. 2015. "Fostering a Culture of Scientific Integrity: Legalistic vs. Scientific Virtue-Based Approaches." *Professional Ethics Report* 2: 1-3.
- Pennock, Robert T. 2018. "Beyond Research Ethics: How Scientific Virtue Theory Reframes and Extends Responsible Conduct of Research." In *Cultivating Moral Character and Virtue in Professional Practices*, edited by David Carr, 166-77. Abingdon, UK: Routledge.
- Pennock, Robert T. 2019. *An Instinct for Truth: Curiosity and the Moral Character of Science*. Cambridge, MA: MIT Press.
- Pennock, Robert T., and Michael O'Rourke. 2017. "Developing a Scientific Virtue-Based Approach to Science Ethics Training." *Science and Engineering Ethics* 23, no. 1: 243-62.
- Show, Nancy, E. 2010. *Virtue as Social Intelligence: An Empirically Grounded Theory*. New York: Routledge.