The Philosophical Use and Misuse of Science

Justine Kingsbury and Tim Dare

Abstract

Science is our best way of finding out about the natural world, and philosophers who write about that world ought to be sensitive to the claims of our best science. However, there are obstacles to outsiders using science well. We think philosophers are prone to misuse science: to give undue weight to results that are untested; to highlight favorable and ignore unfavorable data; to give illegitimate weight to the authority of science; to leap from scientific premises to philosophical conclusions without spelling out their relevance; to treat mere resonance between a scientific theory and a philosophical view as empirical evidence for the philosophical view. In this paper we identify and illustrate some of the ways in which philosophers misuse science, explain why these pitfalls are easy to fall into, and conclude with suggestions for avoiding them.

1. Introduction

Philosophers increasingly appeal to scientific theories and data in support of their philosophical claims, and so they should. Science is our best way of finding out about the natural world, and philosophers who write about that world ought to be sensitive to the claims of our best science. However we think philosophers are prone to misuse

science: to give undue weight to results that are untested; to highlight favorable and ignore unfavorable data; to give illegitimate weight to the authority of science; to leap from scientific premises to philosophical conclusions without spelling out their relevance; to treat mere resonance between a scientific theory and a philosophical view as empirical evidence for the philosophical view. While we think that philosophers are prone to make these errors, we do not mean to suggest that other 'science consumers' – including scientists themselves – are immune: philosophers are prone to cherry-picking, but there are plenty of cherry-picking scientists too. We will argue, however, not only that philosophers are prone to these errors, but also that there is reason for philosophers to be particularly sensitive to the dangers they pose. We conclude with some suggestions about steps that might be taken to minimize the risk.

2. Philosophical misuses of science

We begin with some examples of what we suggest are philosophical misuses of science.

2.1. Cherry-picking

To cherry-pick is to appeal to just those scientific results or theories that appear to support your view and disregard or give insufficient weight to those that conflict with it. Consider John Doris's application of empirical results to philosophical questions about virtue and character (Doris 2002). Doris is aware of the methodological issues that arise when using empirical results, and he cares about using them properly.[1] He relies, he says, "on experiments of the "textbook variety": those representing settled findings that have been replicated and not controverted."

> While there are in some instances conflicting data, I employ, with very few
>
> exceptions… no study for which I have not identified replications or extensions.

2

In making my case, I've tried to limit discussion to experiments that conform to an established pattern of results or have been subjected to substantial critical scrutiny. (Doris 2002, 30, note 3)

Still, Doris does cherry-pick. He cites Isen and Levin's "dime helping experiment" (Isen and Levin 1972) in support of the conclusion that whether or not a person helps someone in trouble is better explained by situation – by transitory and recent influences on mood – than by stable character traits (Doris 2002, 30–32). Isen and Levin planted a dime in the return slot of a pay-phone for some subjects to find; other subjects found nothing in the slot. Once a subject left the telephone booth, a confederate of the researcher dropped a manila folder full of papers in the subject's path. Fourteen of the sixteen people who found a dime helped gather the papers; only one of the twenty-five people who did not find a dime did so. In a variation that removed the effect of a person in need, subjects found an apparently lost letter to mail or not (Levin and Isen 1975). Ten of eleven people who found a dime mailed the letter; only four of thirteen people who did not find a dime did so.

Isen and Levin's results seem philosophically interesting and promising, and Doris is addressing just the sort of question on which empirical research should be brought to bear. Philosophers puzzling about what influences our behavior *should* seek scientific data, rather than relying upon armchair opinions that are likely to be based on anecdotal evidence and personal experience seen through the lens of (often) unacknowledged and unexamined theories.

Primarily because of Doris's use of the dime helping experiment, the story so far will be familiar to many philosophers. Much less widely known is that attempts to replicate Isen and Levin's results have mostly been unsuccessful.[2] Blevins and Murphy re-ran the first dime-finding experiment using dropped packages rather than

dropped papers. Of fifteen people who found a dime, only six helped. Of thirty-five people who didn't find a dime, fifteen helped: "[T]here is", they concluded "no relationship between finding a dime and helping" (1974, p. 326). Weyant and Clark (1977, 109) drew the same conclusion using five different locations and over four times as many subjects: "[S]ubjects who found a dime" they report, "did not mail an apparently lost letter more often than subjects who did not find a dime."[3] In a different study testing the effect of a small piece of good luck on helping behavior, Schellenberg and Blevins (1973) found that people who received an unexpected free certificate for a 'whopper' hamburger were no more likely to volunteer for an experiment than were those who did not. Weyant and Clark point out, referring to their own findings and those of Blevins and Murphy and Schellenburg and Blevins, that these "are not isolated failures to replicate the effect of feeling good on helping behavior in a field study.… The results question the generality and stability of the dime-induced helping phenomena" (Weyant and Clark 1977, 110).

Doris acknowledges the studies by Blevins and Murphy and Schellenberg and Blevins in a footnote, but still, he uses the Isen and Levin study both as evidence for and as a striking illustration of his claim about the influence of situational factors on behavior. He justifies this by saying that in each of the other studies there are departures from the original study that may explain the failure to replicate the results, and that in any event Isen and Levin's experiment "exemplifies an established pattern of results" (Doris 2002, 30, note 4). "Numerous studies have shown that mood can have powerful impacts on a wide variety of human functioning.… Most relevantly, positive affect has repeatedly been shown to be related to prosocial behavior" (Doris 2002, 30). But if there is an established pattern of repeatable results showing that positive affect is related to prosocial behavior, Doris should have cited some of those

results, rather than Isen and Levin, in support of his claims.[4] As it stands, Doris's reliance on Isen and Levin and scant regard to countervailing studies looks like cherry-picking: selecting just those scientific results which support a theory and failing to give proper weight to those which tell against it.[5]

2.2. Using contentious scientific results without appropriate reservation

"Sometimes," writes Greg Cooper (1998, 195), "extreme ideas from one field can show up as uncontroversial truths in another."[6] Using contentious scientific results without appropriate reservation is strikingly illustrated by the enthusiastic reception of mirror neuron research by philosophers working across a wide range of sub-disciplines. Mirror neurons were first observed in macaques in the 1990s. They are so called because the neuron that fires when a macaque performs a task (such as grasping something) also fires when it observes another macaque doing so. Scientists have hypothesized that in humans mirror neurons are the neurological basis of action recognition (Arbib and Rizzolatti 1997; Gallese et al. 1996; Rizzolatti et al. 1996), imitation (Carr et al. 2003), understanding of intentions (Iacoboni et al. 2005), empathy (Rizzolatti and Craighero 2004) and linguistic capacities (Rizzolatti and Arbib 1998), and that mirror neuron deficits may be part of the explanation of autism (Gallagher 2005). However mirror neuron research, particularly on humans, is still at an early stage. Firings of individual mirror neurons have been observed in macaques in numerous studies, but there is so far only one study that detects them in humans (Mukamel et al. 2010). The macaque researchers implanted electrodes in the macaques' brains; most studies of mirror neurons in humans use much less fine-grained functional Magnetic Resonance Imaging (fMRI). A review article that is generally positive about mirror neuron research (Kilner and Lemon 2013) says that the results of fMRI experiments "confirm a broad overlap between cortical areas

active in humans during action observation and areas where mirror neurons have been reported in macaque monkeys," and "seem to be consistent with the existence of a mirror neuron system in humans but they cannot yet furnish conclusive proof." Nonetheless philosophers have appealed to mirror neuron research to support claims about the phenomenology of inter-subjectivity (Dieter 2006), our knowledge of other minds (Goldman 2002), why we perceive music as expressive (Cochrane 2010), and why we emote in response to fictions (Stump 2010) with very little acknowledgement of how speculative the science is.[7]

2.3. Arguing illegitimately from scientific authority

Sometimes when philosophers use scientific theories or results in support of philosophical conclusions, the very fact that those theories are scientific seems to be supposed to give them extra credibility. We might have the intuition, and anecdotal evidence, that having something nice happen to you makes you more generous than you would otherwise be, but Doris has empirical data: 'studies have shown'. This is a kind of rhetorical appeal to the authority of science. We take it that those who make such appeals are not in general thinking that the fact that scientists claim X *makes X true*, as, in law, the fact that the Supreme Court says that something is law *makes it the case* that that thing is law. This kind of constructivism would be inimical to the kind of naturalistic project that philosophers who appeal to science are generally engaged in. Rather, we suppose that they are appealing to the authority of science on the grounds that scientific methods are our best way of finding out about the world: scientific claims, the thought is, are likely to be true.

We might accept this when the scientific claims in question have survived peer scrutiny and testing. However, the mere fact that a study has produced certain results, or that certain results have been published in a reputable scientific journal,

does not give them this kind of authority. The plausible thought that it is in general better to appeal to scientific results than to theorize from the armchair is mistaken if it amounts to the thought that any science will be better than any *a priori* theorizing: bad science, or underdeveloped science, or science that has not yet stood the test of the rigorous processes by which scientific consensus is achieved, is not better than no science at all.

2.4. Applying scientific theories or data too far from their domain: Stump, mirror neurons and the paradox of fiction

Philosophers occasionally misuse scientific results and theories by applying them in domains too far removed from the scientific context in which they were generated. Mirror neurons again provide an illustration, this time used as part of a solution to the paradox of fiction. Why do we feel happy when Elizabeth and Mr. Darcy get together at last, when we know there are no such people? Eleonore Stump thinks that "consideration of the mirror neuron system lets us think about audience appropriation of fiction in a helpful way" (Stump 2010, Chapter 4). There can be no objection to using contentious science to prompt new ways of thinking about old problems, but Stump seems to give mirror neuron research a more substantive role, seeing in that research a specific solution to the problem of fiction, and a solution, furthermore, which depends not only on accepting claimed results from that research but on stretching those results beyond anything in the science itself. "Studies… have shown", writes Stump, "that those parts of the visual system which are involved in the sight of the rotation of objects are also the parts of the system which are used when a person imagines the rotation of imagined objects", and now, she suggests, "[n]othing keeps us from supposing that the mirror neuron system which subserves the knowledge of persons can also be used in this dual purpose way, for the appropriation

of second-person experience either in actuality or in thought only". If this is right, she concludes:

> … then it might be that when we engage with fiction, we also employ the mirror neuron system, but in an alternate mode, just as the visual system is employed in an alternate mode when we imagine the rotation of an imagined object. If the mirror neuron system is like the visual system in this regard, then the same system which explains our knowledge of persons in second-person experience could also explain our appropriation of knowledge of persons through fiction…. When fiction functions as a second-person account and we gain some knowledge of persons from fiction, one possible explanation for why we do so is that the mirror neuron system can also be used in an alternate mode, for the engagement with fiction. (Stump 2010, Chapter 4.)

There are problems with this use of mirror neuron research that are independent of the underdeveloped state of mirror neuron research in humans itself, discussed earlier. Even if the human mirror neuron system is like that for which there is some evidence in macaques, and even if it operates in alternative modes, and even if it were shown that mirror neurons account for our capacity for empathy, we would still be some way from a solution to the problem of emotional responses to fiction. The mirror neuron solution to the problem of empathy would leave unexplained why such neurons were triggered in the absence of actual people experiencing the emotion empathetically experienced, and so it is not obvious how experiments which might support claims about mirror neurons and empathy support claims about mirror neurons and responses to fiction. The grounds for confidence in the original claim about mirror neurons and empathy need to be spelled out, and so does the justification for stretching that claim to cover responses to fiction. As it stands, the use to which

mirror neuron research is being put is a very long way from the questions addressed by the existing research. Stump's remarks are conditional and replete with qualifiers: 'might', 'could', 'possibly'. Nonetheless those remarks are more appropriate as a description of a possible research project than as conclusions about the basis of emotional responses to fiction.

2.5. Drawing philosophical conclusions too quickly from scientific theories or data: Harman on character traits

Sometimes scientific results that resonate with or are suggestive of a philosophical claim are too readily taken as evidence for that philosophical claim. Consider Gilbert Harman's (1999) use of the Good Samaritan experiment. Harman, like Doris, appeals to empirical results that suggest that situational factors have more of an influence on behavior than one might think,[8] but Harman marshals them in support of a much more extreme scepticism about character traits than Doris's.[9]

In the Good Samaritan experiment (Darley and Batson, 1973), students at the Princeton Theological Seminary were asked to walk to another building and give a lecture on a pre-assigned topic. Some were told to hurry and some not, and some were to lecture on the topic of jobs in which seminary students would be most effective and others on the parable of the Good Samaritan. Along the way, they came upon a man slumped against a wall. The topic of the subjects' talks made less difference than one might have expected to their tendency to help the person slumped against the wall. But the degree to which they had been told to hurry made a surprising difference: only 10% of those who were told to rush stopped to help the man, whereas 45% of those who were in a moderate hurry stopped, and 63% of those in no hurry did. Harman takes the experiment to show that people's behavior can be explained by situational factors (such as how much of a hurry they are in), and that consequently there is no

need to appeal to their character traits (such as compassion or the lack of it) to explain their behavior. He concludes that there is no evidence for the existence of character traits.

> Empirical studies designed to test whether people behave differently in ways that might reflect their having different character traits have failed to find relevant differences. It is true that studies of this sort are very difficult to carry out and there have been very few such studies. Nevertheless, the existing studies have had negative results. Since it is possible to explain our ordinary belief in character traits as deriving from certain illusions, we must conclude that there is no empirical basis for the existence of character traits. (Harman 1999, 316)

Responding to subsequent critics Harman goes even further, asserting that there are no character traits – at least, not as ordinarily conceived (Harman 2000, 223-224). As Harman points out, if this is right, it undermines virtue ethics and has important consequences for how we think about moral education: "If there is no such thing as character, then there is no such thing as character building" (Harman 1999, 328).

However, the data leave plenty of room to attribute character traits to the experimental subjects. Arguably, the 10% of those who stopped to help even when in a tearing hurry have the character trait of being compassionate even when under pressure – a trait not possessed by the 90% who didn't help. The experiment (and the Milgram experiment too) might show that people's character traits are less *positive* than might have been expected – fewer people are compassionate under pressure than one would have thought – rather than that people do not have character traits (Athanassoulis 2000, 219). In fact, the experiment seems particularly ill-suited to

settling the role of character, since it does not test the same subjects in different situations. If the same subject tended to help those in the same degree of need in some situations and not others, we might have reason to think it was the situation rather than the agent's dispositions that were doing the work. As it stands, even if the data are good and the results repeatable, Harman's conclusion that there are no character traits is too quick.

3. Why do philosophers misuse science and why does it matter?

We have offered examples of philosophers misusing science, but of course we acknowledge that others, including scientists themselves, do so as well. Enduring (though by no means unanimous) suspicion that Mendel's pea results were too good to be true, Pons and Fleischmann's precipitous announcement of cold fusion, and Andrew Wakefield's misrepresentation of the sample and significance of his MMR study (Wakefield et al. 1998) are familiar examples ranging from alleged cherry-picking to downright scientific fraud. However, we think philosophers appealing to scientific results and theories are at least as likely to misuse science as scientists are if not more so: there are features of philosophy (though they are not unique to philosophy) that facilitate the misuse of science, and features of scientific practice that facilitate the misuse of scientific results by outsiders. We think that the philosophical misuse of science is particularly liable to go undetected, and also that the misuse of science by philosophers is particularly problematic because carefully evaluating methods of inquiry (one's own and others) is part of the philosopher's job.

In almost all cases, a philosopher appealing to scientific theories or scientific data is using the resources of a discipline not her own, with respect to which she is an outsider, and the way science is practiced makes it particularly difficult for outsiders to tell which results and theories can be relied upon. Science done well is a collective

self-correcting enterprise. Some of the correction occurs before publication, both informally and in the peer review process, but much occurs after. Once a scientific paper is published, others see whether the results can be reproduced, conduct modified experiments to test whether the results show what they are alleged to show, and propose alternative hypotheses to explain the results. It may take some time for a consensus to emerge. This is why it is a mistake to think that because a theory or a set of results has appeared in a reputable scientific journal, the theory is true or the results reliable. The process by which scientific knowledge is attained is such that many scientific claims in the public domain will not end up being accepted. [10]

Finding out the result of a particular experiment, or the conclusions researchers have drawn from it, tells us little or nothing without knowledge of the scientific context. Are the results you want to appeal to widely accepted? Are there equally sound studies whose results conflict with them? Is the methodology used one that scientists in the field take seriously? Scientific consensus is often hard to pin down even for those within the field: there are few benchmark acknowledgements of a consensus and fewer still arguments for it. It is easy to see how one might mistakenly give weight to scientific results that do not deserve it.

Other factors that may contribute to the philosophical misuse of science arise not from how science is done but from how philosophy is done. In theory, philosophy is a communal discipline: arguments are published, people respond, we stand on the shoulders of others, making small changes and improvements to their theories and their arguments. In practice, however, the discipline rewards those who present novel or counterintuitive hypotheses: you will do better as a philosopher if you can show that common intuitions about justice or property or knowledge are wrong than you will by saying that everything is just as it seems.[11] The discipline also rewards

individual achievement. Philosophers generally work on their own rather than in teams. Only a small proportion of philosophical work is co-authored, and philosophical works with more than two authors are rare indeed. Most significantly, philosophers tend to see the ideas they defend as *theirs*. It is common for a philosopher to identify herself as a proponent of a particular theory, to set out to defend it, to have made a career out of doing so.[12] This may make it particularly tempting to uncritically appeal to scientific results that look as though they tell in favor of your theory. A partisan defender (particularly in an environment in which it is important to publish often and in which time is at a premium) may seize more quickly upon results that seem to help her case than would a more neutral inquirer.

Furthermore, the nature of the philosophical audience may reduce the likelihood that misuse of science by philosophers will be detected. Some of the readers of empirically informed philosophy will have a good understanding of both scientific methods in general and the particular science at issue, but many will not. Philosophers using science are more likely than scientists to be speaking to a scientifically naïve audience. Philosophers as an audience might be rather easy to impress with apparently spectacular results that seem to provide solutions to long-standing philosophical puzzles.

Finally, one might be particularly concerned about the philosophical misuse of science because such misuse seems strikingly at odds with philosophical ideals. Even if philosophers are no worse than anyone else as regards the misuse of science, one might think that they ought to be held to a higher standard. Philosophers have no excuse for not being reflective about their own epistemic practices. There is also a philosophical ideal of open-minded responsiveness to evidence and argument that sits uncomfortably with many of the practices which amount to misuse of science. In this

respect one might draw a contrast between philosophy and science, on the one hand, and law on the other. Law is openly and explicitly adversarial, and has clear institutional structures recognizing and acknowledging the legitimacy of partisanship (Dare 2009), while philosophy and science are committed, both in theory and as a matter of institutional design, to an ideal of neutrality (Dare and Kingsbury 2008).

4. Some suggestions

We are not suggesting that philosophers should not use scientific theories or scientific data, merely that they should use them with care, and we now make some suggestions (with philosophical illustrations) as to what that might involve. Sections 4.1 – 4.3 suggest ways to avoid cherry-picking, the illegitimate use of marginal science, and illegitimate appeals to scientific authority, and furthermore to make it very clear that these mistakes are being avoided. We acknowledge, however, that it may be extremely difficult for an outsider to judge which scientific results and theories can be relied on, short of completely immersing themselves in the scientific field in question (which is to say, ceasing to be an outsider). Sections 4.4 – 4.5 suggest ways to avoid applying scientific results too far from their original domain; Section 4.6 suggests an institutional change that might make it more difficult for philosophical misuses of science to go unnoticed.

4.1. Being explicit about the role the scientific theory or empirical data is playing in a philosophical argument

There are different ways in which a philosopher might use a scientific theory. She might use it as a premise in an argument, in which case she had better provide good reason to think it is true; this is the kind of case we have mostly been concerned with here. But she might instead use it as an illustration of the possibility of something. Rather than asserting that humans have mirror neurons which are the basis of

empathy, for instance, and that this fact shows that the simulation theory is true, a philosopher might point out that mirror neuron theory describes a way in which humans might, just possibly, find out what's going on in other people's heads by running a simulation. Or a scientific theory might be used conditionally, as in "If mirror neuron theory is true, then the paradox of fiction dissolves."[13] Or a scientific theory might simply spark a new philosophical idea. In the last case, there is no reason to refer to the scientific theory at all; the etiology of the idea plays no role in justifying it. But in the other cases, it would be useful to be very explicit about the role that the scientific theory or set of data is playing. If it is merely illustrating a possibility, for example, then the objection that the scientific theory is underdeveloped or that the results have been cherry-picked will not stick: these are only good objections if the scientific theory or results are being used as premises in an argument.

In 'Ethics and Intuitions' (2005), Peter Singer considers the role moral intuitions have traditionally played in normative ethics, and whether or not recent empirical research (for example Greene and Haidt 2002; Greene et. al. 2001) might show that in fact moral intuitions are ill-suited to that role. Singer is explicit about the status of the research that he appeals to, and about what sort of philosophical use of the research is legitimate given that status: he makes it clear that he is not asserting that the results are correct and then drawing conclusions from them. "These are preliminary results," he says, "based on a limited amount of data. But let us assume that they are sound, and speculate on what might follow from them" (Singer 2005, 342). This is not merely a passing disclaimer that a reader might overlook: cautiously conditional locutions such as "If Greene's initial results are confirmed by subsequent research…" (347) recur throughout.

4.2. Explicitly addressing the credibility of the scientific data or theory

If a scientific theory or set of data are being used as premises, it is important to be explicit about the status of the science. Are the results robust? Are the studies well-designed? Are the conclusions that the researchers draw from the studies justified by the results? Is there a degree of consensus about the theory within the relevant scientific discipline? It would be unrealistic to expect a philosopher to master the methodology of all relevant scientific disciplines. However, it does not seem unreasonable to expect (for example) an awareness of issues concerning sample size and sampling techniques from a philosopher appealing to social science research. Even in appeals to neuroscience, some understanding of the methods by which results are produced may shed considerable light on what conclusions can be drawn from them. Appreciation that the observation of mirror neurons (as opposed to areas of the brain that may contain mirror neurons) requires the insertion of electrodes into the brain, for instance, might raise suspicions that claims about the activity of mirror neurons in human babies are unlikely to be straightforward reports of observed phenomena.

Even where it is not possible for a philosopher to judge what weight a particular set of scientific results will bear by familiarizing herself with the methodology by which those results were produced, it is possible to get some idea by investigating how those results are regarded by others in the field and how they fit with other results. John Doris (2002) (criticized above for cherry-picking) is in general careful about acknowledging weaknesses in scientific research even in cases in which the research tells in his favor. Even in the example we criticize above, Doris acknowledges the existence of countervailing results: our complaint is that he does not give them proper weight. Elsewhere he discusses results which favor his case and

says he cannot base an argument on them because the results are under-described (Doris 2002, 90, note 49); elsewhere again, he points out that there is so much scientific dispute about a particular issue in personality psychology that it would be unreasonable for outsiders to be confident of the truth of either of the competing views ("If the native culture is in such disarray, what's a tourist to think?" (Doris 2002, 73)).

Even in cases in which scientific results are only being used illustratively, it is still a good idea to be explicit about the status of the science you are using as well as about the use to which you are putting it (as Peter Singer is in the case discussed above): doing so may head off criticism.

4.3. Using review articles and meta-studies?

The upshot of our discussion of how science is done is that it is never going to be a good idea to simply read a single original scientific paper and, without further investigation, use its data or its conclusions as a premise in a philosophical argument. Neither is it a good idea to rely on books written by scientists for the general public: they are not subject to such rigorous peer review, and they often over-emphasize the possible connections of their work to issues of general interest without taking care to point out which parts are speculative. (Getting one's science from the mass media – see the discussion of mirror neurons in 2.2 above - is even worse.) When appealing to a particular theory or set of results, it is necessary to find out how that theory or set of results is situated within its discipline.

Consulting review articles and meta-analyses published in well-regarded journals may be a good starting point. However, meta-analyses have their own problems. They may gloss over directly contradictory results, and as Rebecca Jordan-Young points out, studies that appear to be measuring the same thing may not be:

"shifting definitions and measures mean that studies don't actually fit together in the way that they are typically understood to" (Jordan-Young 2010, 200). Hence referring to a meta-analysis does not do away with the need for an informed look at the studies that it aggregates. There are no reliable short-cuts here.

4.4. Collaborating with scientists

Most philosophers do not have the expertise to conduct their own experiments. However, the problem of extrapolating too far from the original field in which results have been gathered could be addressed by directing experiments more precisely at questions of interest to philosophers. For example, the question of what is going on when we respond emotionally to fictions has generated a huge philosophical literature which only occasionally, and not very effectively, appeals to empirical data, partly because there is not much data out there which is directly relevant. However, many of the questions in this area that are of interest to philosophers seem to be empirical questions. What sorts of correlations are there between people's self-reports of emotional responses to fiction and their physiological responses? How do people's physiological responses to a filmed story differ from their physiological responses to reading the same story in a book? What differences of response are there between someone presented with a sad story as a fiction and someone presented with the same story as a documentary? These are not necessarily questions that a psychologist will think to ask, or that a philosopher will have the resources to answer. A philosopher/psychologist collaboration may be a good solution. A further advantage of philosopher/scientist collaborations is that they have the potential to avoid the problem of unfamiliarity with a particular research method and consequent inability to properly evaluate what that method can be used to show.

4.5. Experimental philosophy?

Experimental philosophers are those who do empirical research themselves rather than appealing to other people's results. This avoids one of the problems discussed above – the problem of applying research too far from the domain in which it is gathered – since the experimental philosopher can tailor her research to the philosophical question to be answered. But experimental philosophy is just as liable to the other problems as any other use of empirical data, and possibly more so. For example, someone might be more inclined to rely on a single set of results rather than investigating whether there are others that contradict it if the results in question are her own. There is a further worry about the competence of the researchers: philosophers are not in general trained in empirical methods, and there are numerous examples of flawed study design in experimental philosophy.

Kenneth Shields (2012) points out one recurrent flaw that will serve as an example. A standard technique in experimental philosophy is to elicit the responses of participants to a case described in a paragraph or two (an "intuition probe"). For example, Weinberg, Nichols and Stich present the following Gettier case to their participants:

> "Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been stolen, and he is also not aware that Jill has replaced it with a Pontiac, which is a different kind of American car."

They then ask the participants: "Does Bob really know that Jill drives an American car, or does he only believe it?" (Weinberg, Stich and Nichols (2001), 29). Shields points out that no effort is made in this or in *any* published experimental philosophical study (at the time at which he was writing) to ascertain whether or not

the participants understand the basic facts about the event or situation about which they are being asked. If a participant has not registered (even though it is clearly stated) the fact that in the story above, Jill's current car is a Pontiac, then that participant's answer to the question about Bob's state of knowledge is not going to tell us anything interesting about knowledge. It seems clear that the responses of participants who fail a basic reading comprehension test on the story should not be taken into consideration, and that failure to test participants' reading comprehension is a major flaw in any study which draws philosophical conclusions from participants' responses to intuition probes like the one above. This is just one example of a recurrent flaw in experimental design.

4.6. Rigorous and informed peer review

The suggestions made so far are for individual researchers, but there are also institutional changes that could help to reduce the frequency of philosophical misuse of science. Journals publishing empirically informed papers in philosophy could seek advice from scientists working in the areas in question (not by using them as referees, since they will not be well-placed to judge the philosophical merit of a paper, but by sending excerpts to the relevant experts in additional to normal refereeing processes). Referees themselves should look at the studies cited rather than taking it on trust that the author is representing them accurately. Implementing these suggestions would of course require considerable extra effort on the part of both editors and referees.

5. Conclusion

Philosophers increasingly appeal to scientific theories and data in support of philosophical claims. We have discussed examples of a range of mistakes in the philosophical use of science, pointed out some features of scientific and philosophical

practice that may explain why philosophers make them, and suggested some steps that can be taken to avoid them. Research in other disciplines can profitably be brought to bear on philosophical issues, and we hope to have provided some guidance about how (and how not) to use it.[14]

Justine Kingsbury   Tim Dare

Philosophy Programme   Philosophy Programme

University of Waikato   University of Auckland

Private Bag 3105   Private Bag 92019

Hamilton 3216   Auckland 1142

New Zealand   New Zealand

justine.kingsbury@waikato.ac.nz   t.dare@auckland.ac.nz

References

Arbib, Michael, and G. Rizzolatti. 1997. "Neural expectations: A possible evolutionary path from manual skills to language." *Communication and Cognition* 29: 393-424.

Athanassoulis, Nafsika. 2000. "A Response to Harman: Virtue Ethics and Character Traits." *Proceedings of the Aristotelian Society* 100: 215-21.

Blevins, Gregory, and T. Murphy. 1974. "Feeling good and helping: Further phonebooth findings." *Psychological Reports* 34: 326.

Carr, Laurie, et. al. 2003. "Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas." *PNAS* 100: 5497-502.

Cochrane, Tom. 2010. "A Simulation Theory of Musical Expressiveness." *Australasian Journal of Philosophy* 88, no. 2: 195-207.

Cooper, Gregory. 1998. "Teleology and Environmental Ethics." *American Philosophical Quarterly* 35, no. 2: 195-207.

Dare, Tim, and J. Kingsbury. 2008. "Putting the Burden of Proof in its Place: When Are Differential Allocations Legitimate?" *Southern Journal of Philosophy* 46: 503–18.

Dare, Tim. 2009. *The Counsel of Rogues? A Defence of the Standard Conception of the Lawyer's Role*. London: Routledge.

Darley, John, and C.D. Batson. 1973. "'From Jerusalem to Jericho': A Study of Situational and Dispositional Variables in Helping Behavior." *Journal of Personality and Social Psychology* 27: 100 -108.

Doris, John. 2005. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.

Gallagher, Shaun. 2005. *How the Body Shapes the Mind*. Oxford: Clarendon Press.

Gallese, Vittorio, et al. 1996. "Action recognition in the premotor cortex." *Brain* 119: 593-609.

Goldman, Alvin. 2002. "Simulation Theory and Mental Concepts." In *Simulation and Knowledge of Action,* edited by Jerome Dokic and Joelle Proust, 1–19. Amsterdam: John Benjamins.

Gopnik, Alison. 2007. "Cells That Read Minds? What the Myth of Mirror Neurons Gets Wrong About the Human Brain." *Slate*: Special Issue on Neuroscience and Neuroculture.

Greene, Joshua, R.B. Somerville, L.E. Nystrom, J.M. Darley, and J.D. Cohen. 2001. "An fMRI Investigation of Emotional Engagement in Moral Judgment." *Science* 293: 2105–8.

Greene, Joshua, and J. Haidt. 2002. "How (and Where) Does Moral Judgment Work." *Trends in Cognitive Sciences* 6: 517-23.

Harman, Gilbert. 1999. "Moral Psychology Meets Social Philosophy: Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99: 315-31.

Harman, Gilbert. 2000. "The Non-existence of Character Traits." *Proceedings of the Aristotelian Society* 100: 223-26.

Iacoboni, Marco, I. Molnar-Szakacs, V. Gallese, G. Buccino, J.C. Mazziotta, and G. Rizzolatti. 2005. "Grasping the intentions of others with one's own mirror neuron system." *PloS Biology* 3: 529-35.

Isen, Alice, and P. Levin. 1972. "Effect of Feeling Good on Helping: Cookies and Kindness." *Journal of Personality and Social Psychology* 21:384-88.

Kilner, J., and R. Lemon. 2013. "What We Know Currently About Mirror Neurons." *Current Biology* 23: R1057-62.

Levin, Paula, and A. Isen. 1975. "Further studies of the effect of feeling good on helping." *Sociometry* 38:141–47.

Jordan-Young, Rebecca. 2010. *Brainstorm: The Flaws in the Science of Sex Differences*. Cambridge, Mass.: Harvard University Press.

Lohmar, Dieter. 2006. "Mirror Neurons and the Phenomenology of Intersubjectivity." *Phenomenology and the Cognitive Sciences* 5: 5-16.

Milgram, Stanley. 1963. "Behavioral Study of Obedience." *Journal of Abnormal and Social Psychology* 67:371–78.

Miller, Christian. 2009. 'Social Psychology, Mood, and Helping: Mixed Results for Virtue Ethics." *Journal of Ethics* 13:145-73

Mukamel, Roy, et al. 2010. "Single-neuron responses in humans during execution and observation of actions." *Current Biology* 20:750-56.

Rizzolatti, Giacomo, et al. 1996. "Premotor cortex and the recognition of motor actions." *Cognitive Brain Research* 3: 131-41.

Rizzolatti, Giacomo, and M. Arbib. 1998. "Language within our grasp." *Trends in Neurosciences* 21:188-94.

Rizzolatti, Giacomo, and L. Craighero. 2004. "The Mirror Neuron System." *Annual Review of Neuroscience* 27:169-92.

Ross, Lee, and Richard Nisbett. 1991. *The Person and the Situation: Perspectives of Social Psychology.* Philadelphia: Temple University Press.

Schellenberg, James, and G. Blevins. 1973. "Feeling good and helping: How quickly does the smile of Dame Fortune fade?" *Psychological Reports* 33:72–74.

Shields, Kenneth. 2012. "Motivational Internalism and Experimental Philosophy." Presented at the Buffalo Experimental Philosophy Conference, Buffalo, NY.

Singer, Peter. 2005. "Ethics and Intuitions." *The Journal of Ethics* 9:331-52.

Stump, Eleonore. 2010. *Wandering in darkness: narrative and the problem of suffering.* Oxford: Oxford University Press.

Wakefield, Andrew, et. al. 1998. "Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children." *The Lancet* 351: 637–41. (Retracted)

Weinberg, Jonathan, S. Nichols and S. Stich. 2001. "Normativity and Epistemic Intuitions." *Philosophical Topics* 29:429-60. Reprinted in *Experimental Philosophy*, edited by Joshua Knobe and Shaun Nichols, 17-45. Oxford: Oxford University Press.

Weyant, James, and R. Clark, R. 1977. "Dimes and Helping: The Other Side of the

    Coin." *Personality and Social Psychology Bulletin* 3:107–10.

---

[1] In Section 4 below, we offer Doris as an exemplar of some of our suggestions for careful use of science.

[2] It is not that philosophers have never noticed this problem: see for instance Miller 2009, 148.

[3] Of fifty-four people who found a dime, forty-two failed to help; of fifty-two people who didn't find a dime, thirty-seven failed to help.

[4] Miller makes this point: "To his credit … Doris does acknowledge the replication trouble for Isen and Levin's experiments. Given the wealth of other similar experiments, though, it is not clear why he did not appeal directly to them instead" (Miller 2009, 149).

[5] It is worth noting that the emphasis Doris (and Gilbert Harman 1999 – see 2.4 below) puts on some of these experiments at the expense of others reflects the way in which the same experiments are treated by the social psychologists Lee Ross and Richard Nisbett (1991). Scientists also cherry-pick.

[6] Cooper offers several examples in which environmental ethicists use controversial claims from ecology and ethology as though they were uncontentious.

[7] The enthusiastic embracing of mirror neurons is not limited to philosophers. Alison Gopnik (2007) points out that mirror neurons have captured the popular imagination, and suggests that they are the "right-brain/left-brain of the 21st century". Uncritical acceptance of the way mirror neuron research is portrayed in popular science magazines and other media outlets may partly explain its popularity in philosophy and other academic disciplines.

[8] Harman appeals to the Milgram experiments (Milgram 1963) as well as the Good Samaritan experiment. We will focus on his use of the latter, but we think there are similar problems with his use of the Milgram experiments.

[9] See Doris 2002, 23, note 36. Doris is skeptical only about what he calls "global" character traits, whereas Harman appears to be skeptical about character traits *tout court*.

[10] This is not a criticism of how science works: the process by which results and theories are made publicly available for criticism is precisely what makes it likely that the theories that emerge from that process will be ones we can rely on.

[11] This feature is of course not unique to philosophy. For example, the appeal of Ross and Nisbett's (1991) situationism in social psychology is surely attributable in part to its novelty and counterintuitiveness.

[12] The tendency of researchers to identify as the proponent of a particular view and give a partisan defense of it is not unique to philosophy either.

[13] Stump (quoted in 2.4 above) *does* use mirror neuron research conditionally, but (as we have argued above) there are nevertheless problems with her use of it.

[14] Numerous people have provided helpful feedback on earlier versions of this paper – too many to list. Thanks to all of you, and in particular to Tracy Bowell, David Braddon-Mitchell, Greg Cooper, Maite Ezcurdia, Patrick Girard, Paul Griffiths, Cathy Legg, Susana Lucero, Jonathan McKeown-Green, Glen Pettigrove, Kelly Roe, John Sarnecki, Angela Smith, and an anonymous referee.