

## Plato's Method Meets Cognitive Science

Normative questions – particularly questions about what we should believe and how we should behave – have always been high on the agenda for philosophers, and over the centuries there has been no shortage of answers proposed. But this abundance of answers raises yet another fundamental philosophical question: How should we *evaluate* the proposed answers; how can we determine whether an answer to a normative question is a good one? The best known and most widely used method for evaluating answers to normative questions can be traced all the way back to Plato. Recently, however, cognitive scientists interested in cross cultural differences have reported findings that pose a serious challenge to this venerable philosophical method. Indeed, in light of these new findings some philosophers – I am one of them – have come to think that after 2400 years it may be time for philosophy to stop relying on Plato's method. In the pages that follow I'll sketch the path that led me to this conclusion.

To introduce the method, let's begin with an example of its use by one of its most brilliant practitioners, Plato himself. Here is a famous passage from *The Republic* in which Socrates recounts using the method in a conversation about the nature of justice.

Well said, Cephalus, I replied: but as concerning justice, what is it? – to speak the truth and to pay your debts – no more than this? And even to this are there not exceptions? Suppose a friend when in his right mind has deposited arms with me and he asks for them when he is not in his right mind, ought I to give them back to him? No one would say that I ought or that I should be right in doing so, any more than they would say that I ought always to speak the truth to one who is in his condition.

You are quite right, he replied.

But then, I said, speaking the truth and paying your debts is not a correct definition of justice.

Quite correct, Socrates.<sup>1</sup>

---

<sup>1</sup> *The Dialogues of Plato*, translated by B. Jowett. I, 131, p. 595. New York: Random House

The central idea of Plato's method, clearly on display here, is to test normative claims against people's spontaneous judgments about real and hypothetical cases. Contemporary philosophers often call these spontaneous judgments "intuitions." If the normative claim and people's intuitions agree, the claim is vindicated. But if, as in Socrates' dialogue with Cephalus, a normative principle conflicts with people's intuitions, then something has to give. Sometimes we may hold on to the normative claim and ignore a recalcitrant intuition. But if a normative principle conflicts with lots of intuitions or, as in the example from *The Republic*, if it conflicts with an intuition that we would be very reluctant to give up, then Plato's method requires that we reject the principle and try to come up with another one.

Though philosophers have been using this method for over two millennia, it is far from clear why we should trust it. Why should we think that a normative principle that has been sanctioned by the method is likely to be a good one? Indeed, why do principles that cohere with our intuitions have any special status at all? Plato, of course, had an answer to this question. It was part of his famous theory of *anamnesis* or recollection. Though scholars would give a much more nuanced account, the basic idea is that before we were born our souls had an opportunity to gain knowledge of the Forms which determine the true nature of knowledge, justice, piety – and everything else. When the soul enters the body the whole business is so traumatic that the soul forgets what it knew about the Forms. Fortunately, the knowledge is not totally lost. It still guides our judgments about cases like the one that Socrates poses for Cephalus, and by using Plato's method with diligence we can succeed in recovering explicit knowledge of the Forms.

It's certainly a ingenious story, though even in Plato's time few people accepted it, and it would be hard to find a contemporary philosopher who takes the stuff about the soul's prior encounter with the Forms at all seriously. But useable methods are not exactly thick on the ground in philosophy, so while Plato's account of why his method works has been roundly rejected, the method itself most definitely has not. Consider, for example, the following enormously influential passage in which Nelson Goodman, one of the great analytic philosophers of the last half of the 20<sup>th</sup> century, offers a wonderfully lucid account of the method that he and others in the analytic tradition have long been using.

How do we justify a *deduction*? Plainly by showing that it conforms to the general rules of deductive inference....Analogously, the basic task in justifying an inductive inference is to show that it conforms to the general rules of *induction*.

Yet, of course, the rules themselves must eventually be justified. The validity of a deduction depends not upon conformity to any purely arbitrary rules we may contrive, but upon conformity to valid rules.... But how is the validity of rules to be determined? ...Principles of deductive inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences that we actually make and sanction. If a rule yields unacceptable inferences, we drop it as invalid. Justification of general rules thus derives from judgments rejecting or accepting particular deductive inferences.

This looks flagrantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.* The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either.<sup>2</sup>

In *A Theory of Justice*, one of the most influential philosophical books of the last fifty years, John Rawls advocates much the same method for justifying principles in the moral domain. Rawls also introduced a new name for the method; he called it *the method of reflective equilibrium*.

But what about our concern over the justification of the method. What do contemporary philosophers have on offer to replace Plato's myth about the prenatal adventures of the soul? Goodman, it seems, thinks that passing the reflective equilibrium test is (as philosophers sometimes say) *constitutive* of justification or validity for normative principles. Passing the reflective

---

<sup>2</sup> *Fact, Fiction and Forecast*, pp. 66-67. Indianapolis: Bobbs-Merrill (1965).

equilibrium test is what it *is* for a principle to be justified. There is no need to tell tales about dimly remembered encounters with the Forms. In the agreement between principles and intuitions, Goodman maintains, lies the only justification that is needed for either.

I first encountered Goodman's account of the reflective equilibrium method when, as an undergraduate at the University of Pennsylvania in the early 1960s, I had the great good fortune to take a class with Goodman. It all seemed overwhelmingly plausible to me at that time and for many years after. During those years, I had some very good company, since the method was ubiquitous in post World War II Anglo-American philosophy. But then, about 15 years ago, I began to have some doubts. In my initial attempts to articulate those doubts I did what philosophers so often do – I described a hypothetical case: Suppose we were to encounter cognitive agents – Martians, perhaps, or members of an exotic tribe – who reasoned and formed beliefs very differently from the way we do. Suppose further that these exotic folks also have very different intuitions about reasoning, and that when the method of reflective equilibrium is used with their intuitions it turns out that their way of reasoning is justified, though when the method is used with our intuitions, it turns out that our very different way of reasoning is justified. Surely, I argued, something has gone very wrong here, since it seems that the method of reflective equilibrium can justify *any* way of going about the business of reasoning, no matter how bizarre, so long as the folks who use the method have the intuitions to match. We could, of course, avoid the problem by insisting that in employing the method we must use *our* intuitions, not theirs. But it is (to put it mildly) less than obvious how this move could be defended. Since we don't believe that our intuitions have been shaped by a prenatal glimpse at the Forms, why should we privilege our intuitions over those of some other group?

Though I thought this was a rather clever objection to the method of reflective equilibrium, others were unconvinced. And they had an interesting argument: We rely on *our* intuitions in philosophy for much the same reason that we rely on our intuitions in mathematics. They are the only game in town – the only *real* intuitions that anyone has. Why should we worry about what might be justified by the intuitions of imaginary Martians, they asked, when there are no such Martians? That's a good question, and one to which, until recently, I had no good answer.

About three years ago, however, I happened to run across my old friend, the psychologist Richard Nisbett, and he began to tell me about some

enormously exciting experiments that he has been doing. For decades, scholars have been claiming that people in East Asian cultures have very different “mentalities” from people in Western cultures. The Chinese, scholars have claimed, and others influenced by Chinese culture, perceive and think about the world around them in very different ways from people in Western cultures, and these differences are reflected in the way they describe and explain events and in the beliefs and theories they accept. Nisbett and his colleagues had begun to explore whether these claims about differences in mentalities could be experimentally verified, and – to my amazement and to the amazement of most cognitive scientists as well – they discovered that many of them could. There is a growing body of evidence indicating large and systematic differences between East Asians and Westerners on a long list of cognitive processes including attention, memory and perception. The two groups also differ in the way they go about describing, predicting and explaining events, in the way they categorize objects and in the way they revise beliefs in the face of new arguments and evidence.<sup>3</sup> This work suddenly made it very plausible that the first part of my hypothetical case is more than just a philosophical fantasy. There really are people whose reasoning and belief forming strategies are very different from ours. Indeed, there are over a billion of them!

But what about the second part of my thought experiment, the part that focused on intuitions? Nisbett hadn’t looked for cross cultural differences in intuitions, but when I mentioned the possibility to him, he thought it was worth a try. So, in collaboration with Shaun Nichols and Jonathan Weinberg, I decided to run a few experiments designed to explore whether in addition to the differences that Nisbett found between Asian and Western processes of acquiring knowledge, there might also be differences between Asian and Western intuitions about what knowledge *is*.

One of the oldest and most durable doctrines in epistemology is an account of knowledge that was first proposed by Plato in a dialog called the *Theaetetus*. To count as an instance of knowledge, on this view, a belief must be true, and the believer must have some justification for believing it. Neither lucky guesses nor false beliefs count as knowledge. This analysis of knowledge as *justified true belief* was the received view from Plato’s time until 1963, when Edmund Gettier published a number of hypothetical cases in which people had justified true beliefs, though what made the beliefs true was not causally related

---

<sup>3</sup> R. Nisbett, K. Peng, I. Choi, & A. Norenzayan, “Culture and systems of thought: holistic vs. analytic cognition.” To appear in *Psychological Review*, 2001.

to what made the beliefs justified. Here is an example of the sort of case that Gettier proposed.

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. Does Bob really know that Jill drives an American car, or does he only believe it?

Cases like this have had a vast impact on philosophers. Just about every epistemologist has the strong intuition that Bob does not really have knowledge in “Gettier cases,” and those intuitions have led the overwhelming majority of philosophers to conclude that Plato’s *justified true belief* account of knowledge is mistaken or incomplete.

But Nichols, Weinberg and I had a hunch. Gettier cases are typically very similar to unproblematic cases in which the fact that makes the belief true *is* causally involved in the justification of the belief. And Nisbett’s group has shown that East Asians are more inclined than Westerners to make judgments on the basis of similarity. Westerners, on the other hand, are more disposed to focus on causation in describing the world and classifying things. So, we speculated, perhaps East Asians might be somewhat less inclined than Westerners withhold the attribution of knowledge in Gettier cases. The results of an experiment designed to test this speculation were nothing short of startling. As expected a substantial majority of Western subjects (74%) claim that Bob *only believes* that Jill drives an American car, but a majority of East Asian subjects (57%) claim that Bob *really knows!* With South Asian subjects, the difference is even more remarkable. Sixty one percent of our Indian, Pakistani and Bangladeshi subjects report that Bob *really knows*. The “Gettier intuitions,” that led to the rejection of the Platonic account of knowledge are, it appears, *very* culturally local.

Encouraged by these findings, we embarked on a much more ambitious project designed to test a variety of philosophical “intuition probes” on a variety of different groups. So far we’ve found a total of six intuition probes – all modeled on hypothetical cases that have been widely discussed by epistemologists – on which different groups have significantly different intuitions. In some cases the differences are between people with different cultural backgrounds, while in other cases they are between people of different

socio-economic status (SES). High SES Americans and low SES Americans have different epistemic intuitions! Moreover, in many cases these differences are quite dramatic. So why haven't philosophers noticed them? Well, since years of education is a major factor in determining SES, *all* philosophy professors are high SES, and the overwhelming majority of us are white and of Western European ancestry.

What conclusions should we draw from these studies? Perhaps the first thing to say is that it is early days yet. All six of the intuition probes on which we've found significant group differences ask subjects to say whether the case counts as knowledge or mere belief. It would be fascinating to know whether there are also cultural and SES differences when subjects are asked to judge what a person *should believe* or whether a belief is *justified*, and we are currently collecting data using these intuition probes. Further down the road, we'd like to look at moral intuitions to see if they exhibit the same cultural diversity that we're finding in epistemic intuitions.<sup>4</sup> If we continue to find the sorts of systematic cultural differences in philosophical intuitions that our first studies have uncovered, the conclusion I'd be inclined to draw is that *Plato's method should be rejected as a strategy for answering normative questions*, since it will yield wildly different results depending on whose intuitions we use. Another option would be to go relativistic and conclude that the epistemic and moral norms appropriate for the rich are different from those appropriate for the poor, and that the norms appropriate for white people are different from those appropriate for people of color. But even if one is inclined to relativism in matters normative – and I am – this strikes me as relativism gone mad. Surely a much more reasonable conclusion is that philosophy's 2400 year long infatuation with Plato's method has been a terrible mistake.<sup>5</sup>

Stephen P. Stich

Board of Governors Professor of  
Philosophy & Cognitive Science

Rutgers University

---

<sup>4</sup> There is already some very impressive evidence indicating that people in different SES groups have different moral intuitions. See J. Haidt, S. Koller & M. Dias. "Affect, Culture and Morality," *J. of Personality & Social Psychology*, 65, 4, 613-628, 1993.

<sup>5</sup> I'd like to thank Shaun Nichols for his helpful feedback on an earlier draft of this article.