

## Chapter 7

### Reflections on Reflective Equilibrium

*Robert Cummins*

#### **Reflective Equilibrium and Scientific Method**

As a procedure, reflective equilibrium (RE) is simply a familiar kind of standard scientific method with a new name. (For descriptions of reflective equilibrium, see Daniels 1979, 1980b, 1984; Goodman 1965; Rawls 1971.) A theory is constructed to account for a set of observations. Recalcitrant data may be rejected as noise or explained away as the effects of interference of some sort. Recalcitrant data that cannot be plausibly dismissed force emendations in theory. What counts as a plausible dismissal depends, among other things, on the going theory, as well as on background theory and on knowledge that may be relevant to understanding the experimental design that is generating the observations, including knowledge of the apparatus and observation conditions. This sort of mutual adjustment between theory and data is a familiar feature of scientific practice. Whatever authority RE seems to have comes, I think, from a tacit or explicit recognition that it has the same form as this familiar sort of scientific inference.

One way to see the rationale underlying this procedure in science is to focus on prediction. Think of prediction as a matter of projecting what is known onto uncharted territory. To do this, you need a vehicle—a theory—that captures some invariant or pattern in what is known so that you can project it onto the unknown. How convincing the projection is depends on two factors: (i) how sure one is of the observational base, and (ii) how sure one is that the theory gets the invariants right. The two factors are not independent, of course. One's confidence in the observational base will be affected by how persuasively the theory identifies and dismisses noise; one's confidence in the theory, on the other hand, will depend on one's confidence in the observations it takes seriously. Prediction is important as a test of theory precisely because verified predictions seem to show that the theory has correctly captured the general in the particular, that it has got the drift of the observational evidence in which our confidence is ultimately grounded. Falsified prediction seems to show that it has not. We are justified in accepting a theory to the extent that we are justified in thinking it properly transfers our confidence concerning observed cases to those that have not been observed. Theory is certainly more than a vehicle for sophisticated inductive inference, but it needs to be at least that if it is to count as more than mere speculation.

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

Because RE has the same form as standard theory construction, it seems reasonable to ask whether it has the same rationale. RE, one might reasonably suppose, is epistemologically normative just in case it too can be seen as a sophisticated case of inductive inference. To explore this idea, imagine the application of a set of principles in RE to a case in which intuition is silent. For the result to be authoritative, we must take the principles to constitute a theory of the property P that is the target of the intuitions, a theory that captures whatever it is that makes something a P. In such a case, the theory constitutes a bridge that transfers our relative certainty about the clear cases to an unclear case. In principle, this could be a simple matter of enumerative induction, but in practice, a theory is required because the intuitions do not typically constitute a set of observations over which we can execute a simple enumerative induction. Imagine, for example, a set of judgments about solutions to various distribution problems in which the proposed solutions are judged fair or unfair.

Experiment one:

Several children are playing on the beach, and they uncover a box of pennies: (i) each child receives the same number of pennies; (ii) they are divided according to the relative sizes of the children, with the largest getting most, and so on; (iii) one child divides the pennies into groups, and the others choose in an order determined by drawing lots; (iv) they are divided according to wealth, with the poorest receiving most, and the richest receiving least.

Experiment two:

Several children are playing on the beach, and they uncover a box of mixed jewels: (i) each child receives the same number of jewels; (ii) they are divided according to the sizes of the children; (iii) one child divides them into groups, and the others pick in an order determined by lot; (iv) they are divided according to wealth, with the poorest receiving most, and the richest receiving least.

A set of "fair" or "unfair" judgments cannot be mechanically generalized in the way you can generalize from a sample of swans to the color of all swans (or to the next swan to be observed). But the underlying epistemological goal is the same: to find a bridge that projects what is known onto uncharted territory.

A theory in RE certainly makes predictions about cases on which intuition is silent or unstable. But how is one to tell if these predictions are successful? They cannot be checked against intuition, for, by hypothesis, intuition is silent or unstable on the cases in question.'

But perhaps this gets the analogy wrong. We do not test scientific theory on cases in which observation is impossible or unstable. We rather make predictions about cases in which the relevant observations have not been made, or were not part of the observational base from which the theory was generated. So perhaps we should think of our intuitively based philosophical theories as making predictions about cases on which intuition has not yet been consulted, or on cases that were not part of the intuitive base from which the theory was generated. This seems a pretty fair representation of philosophical practice. Theory is tested against intuitions about cases dreamed up or brought to attention by the opposition, and we can count on these not

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

being cases involved in the generation of the theory. Fair enough: theories in RE are subject to predictive test just as their scientific look-alikes are. If we want to criticize RE, then, we are squarely reduced to asking what authority the intuitions themselves have; to asking, that is, whether intuition can play the role observation plays in science.<sup>2</sup>

## **Intuition and Observation**

We can dismiss out of hand the idea that philosophical intuition is a case of perception, since it is elicited by linguistic specifications of actual and hypothetical cases. In comparing philosophical intuition to observation, then, I am using 'observation' as this term is used in texts on experimental design and statistics, to mean a data point. Understood this way, it makes sense to ask whether philosophical intuitions stand to philosophical theories as observations stand to scientific theories.<sup>3</sup>

We can get this question in clearer focus by asking what should happen when we are faced with incompatible intuitions. Incompatible scientific observations must be dealt with in one of two ways. Either one or both of the observations must be plausibly dismissed as error or artifact, or both must be discounted. To a good first approximation, observations are admissible scientific evidence only if they are intersubjective. Observations that can, in principle, be made by only one party to a dispute do not count as evidence. Of course, anyone can have an intuition about a given case. But this is not what the intersubjectivity of observation requires. It requires rather that when we make the same observations, we get the same results. If I see a clockwise rotation of the constellations when I look at the night stars, but you see a counterclockwise rotation, we must either explain away one or both observations, or we will have to discard both. We can-not, in this instance, agree to disagree, because we cannot maintain the objectivity of science while tolerating a fundamental subjectivity in the observational evidence. Correspondingly, if one of two conflicting intuitions cannot be explained away, both must be discarded as evidence if RE is to yield theories with any intersubjective force. One might reasonably take the view, then, that a theory is not in RE until conflicts of intuition have been resolved. Taking this line concedes that a theory in RE has no power to settle the dispute, for we have ruled that it is not in RE until the dispute is settled. But this should come as no surprise: Genuine disputes about the data cannot be settled by a theory those data are supposed to test. The data can, occasionally, settle disputes about theory, but the theory cannot settle disputes about the data. Perhaps I cannot help what I intuit any more than I can help what I see, but when conflicts cannot be explained away, I must make no evidential appeal to observations or intuitions in conflict with others.

I belabor this point because I think it is little honored in practice. Philosophers do labor to explain away the conflicting intuitions of others, but it is most assuredly not standard practice to put disputed intuitions in escrow, as it were, pending resolution. Theorists regularly claim support from disputed intuitions when there is no resolution in sight. Indeed, disputed intuitions are often the linchpin on which everything turns. Consider the role of Twin-Earth cases in current theories of content. It is commonplace for researchers in the Theory of Content to proceed as if the relevant intuitions were undisputed (Fodor 1990, 1994).

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

Nor is the reason for this practice far to seek. The Putnamian take (Putnam 1975) on these cases is widely enough shared to allow for a range of thriving intramural sports among believers. Those who do not share the intuition are simply not invited to the games. This kind of selection allows things to move forward, but it has its price. Since most nonphilosophers do not share the intuition, the resulting theories of content have little weight with them, and this is surely a drawback for a theory that is supposed to form an essential part of the foundations of cognitive psychology. Making a Putnamian conscience an entrance requirement for the theory of content threatens to make it irrelevant. We must take care that such agreement about the intuitions as there is is not merely a selection effect. This is easier said than done, since it is all too easy for insiders to suppose that dissenters just do not understand the case. If we are honest with ourselves, I think we will have to confront the fact that selection effects like this are likely to be pretty widespread in contemporary philosophy.

The role of Twin-Earth examples in the theory of content is just an example, but it is significant because there are deep reasons why serious effort is seldom devoted to explaining away divergent philosophical intuitions, and relying instead on social pressures to maintain enough uniformity to keep the game going. For, at bottom, *all* philosophical intuition can be explained away. To get to this bottom line, we will need to go a little indirectly. I propose first to explain why philosophical intuition, unlike scientific observation, is never calibrated, though it certainly could be. This discussion will set the stage for an argument designed to show that all philosophical intuitions are likely to be artifacts.

### **Calibration**

Observations are explained away as errors or artifacts. Errors range from such gross mistakes as pointing the telescope in the wrong direction, to subtle matters of accuracy. An artifact is an observation that carries information about the observational apparatus or process rather than about the target. An example is the apparent curvature of objects in peripheral vision introduced by astigmatism or reading glasses.

Every scientific subdiscipline spends a good deal of effort identifying and correcting errors and artifacts. What is important for present purposes is that an observational technique is deemed acceptable just to the extent that it can be relied upon to produce accurate representations or indicators of its targets. This is why observational procedures in general, and instruments in particular, have to be calibrated. When Galileo pointed his newly devised telescope at the moon and saw mountains—earthlike blemishes on what should have been a perfect celestial object—it was legitimate for the opposition to inquire whether the apparent mountains were artifacts. The proper response was to point the telescope at something of known size, shape, distance, color, and so on to determine what distortions it introduced; to calibrate it, in short.

The details of calibration can be subtle and vary considerably from case to case. An invariable requirement, however, is that there be, in at least some cases, access to the target that is independent of the instrument or procedure to be calibrated. That access need not

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

itself be observational, but it does need to be independent in the sense that it cannot rely on either the instrument or procedure being tested, nor on the theory the newly generated observations are recruited to support. Galileo could not legitimately argue, for example, that the mountains were unlikely to be artifacts on the grounds that the Copernican hypothesis he was concerned to support abolishes the distinction between terrestrial and celestial objects that makes lunar mountains an embarrassment.

Unlike genuine observational techniques, no one ever attempts to calibrate philosophical intuition. To see why, consider again the fairness example. Suppose we want to calibrate someone's fairness intuition. What we need is a list of representative distributions, together with a test key, that is, something that tells us which distributions on the list are fair and which unfair, or, perhaps, a ranking. We then let our subject take the test and see how well he does. Trouble is, where do we get the key? We might try to use only uncontroversial cases—that is, cases everyone agrees on in advance. But if we know that everyone, including our subject, agrees on the test items, there is no point in administering the test. If the subject gives a "wrong" answer, that just shows that the item did not belong on the test. Of course, we could simply be looking to see if the subject is like everyone else we have tested so far. This might be interesting. Perhaps, pursuing this strategy, we might find that everyone, or nearly everyone, or nearly everyone in a certain culture, or economic class, or what have you, shares certain fairness judgments. That would be worth knowing, but it would not count as calibrating our subject's fairness intuition.

We could, perhaps, get a key by consulting the experts. This is unobjectionable, provided the experts do not simply consult their fairness intuitions, for, of course, they have not yet been calibrated. Fortunately, what the experts are certain to do is to apply the best theory of fairness they have to the cases on the test. Of course, if there is a lot of controversy about the best theory of fairness, our experts are liable to disagree. If that happens, we will just have to wait until more is known. And even if the experts agree, they might still be wrong. So we would have to qualify our conclusion: "On the assumption that current theories of distributive justice are on the right track," we might report, "the subject appears to have reliable intuitions about fairness." Not definitive, perhaps, but no worse than the sort of conclusions we can draw about electron microscopes. A final and crucial proviso is that the experts do not base their theories on intuition, for this would evidently launch us on a regress.'

So philosophical intuition could be calibrated, but only on the assumption that there is some nonintuitive access to its targets. Personally, I am inclined to think there are, at least in some cases, nonintuitive routes to the targets of philosophical intuition. We can give up on intuitions about the nature of space and time and ask instead what sort of beasts space and time must be if current physical theory is to be true and explanatory. We can give up on intuitions about representational content and ask instead what representation must be if current cognitive theory is to be true and explanatory. But even if philosophical intuition *can* be calibrated, it never *is* calibrated, because philosophers could have no possible use for intuition in a context in which the relevant theory was well enough settled to form the basis of a credible calibration test. Philosophical theory in such good shape is ready to bid the Socratic midwife farewell and strike out on its own in some other department. Philosophical intuition, therefore, is epistemologically useless, since it can be calibrated only when it is not

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

needed. Once we are in a position to identify artifacts and errors in intuition, philosophy no longer has any use for it. But if we are *not* in a position to do this, philosophy should not have any faith in it.

### **The Probable Sources of Philosophical Intuition**

In spite of what I have just argued, I think we can be pretty confident that philosophical intuitions *are* artifacts. *Consider the source*, as we old folks used to say when doubting press releases from the U.S. military in VietNam. Given the current state of knowledge, I think there are just the following possible sources of philosophical intuitions.

1. *Explicit Theory*. Philosophical intuitions are the result of applying some more or less explicit theory.
2. *Ordinary Beliefs*. Philosophical intuitions are, or are obvious consequences of, things people simply happen to believe as the result of education, socialization, and the like.
3. *Language*. Philosophical intuitions are, or are generated by, the knowledge base one acquires in acquiring one's language.
4. *Concepts*. Philosophical intuitions are generated by concepts.
5. *Tacit Theory*. Philosophical intuitions are the result of applying some tacit theory.

I propose to look at each of these in turn. Some will quickly reduce to others, so the survey will be more manageable than one might suppose.

### **Explicit Theory**

Imagine two cannon balls of exactly the same size—one solid, one hollow—dropped simultaneously from a height of 100 feet. Which will hit the ground first? My "considered judgment" is that they will land at the same time. Some people think the solid ball will land first. No one thinks the hollow ball will land first.

There is no question about the origin of my considered judgment: I can remember when I learned the relevant theory and its application to this case in high school physics. I remember the name of my teacher, how he looked, and how the classroom looked. I remember all this, because this particular application of the theory made a profound impression on me. It had this effect because, prior to that time, I would certainly have said that the solid ball would land first, a belief I picked up partly from my own experience and partly from the implications of things others said.

There is, of course, nothing whatever wrong with "considered judgments" that are applications of an explicit theory. Indeed, it is one of the important functions of an explicit theory to generate such judgments. But the judgments, or intuitions, as we have been calling

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

them, cannot be cited as evidence for the theories that generate them. Those "intuitions" might, of course, be cited as evidence for some other theory, but their epistemological status then evidently reduces to that of the theory that generates them. They have no epistemological weight of their own. In general, intuitive judgment generated by explicit theory inherits whatever epistemological weight belongs to its parent. I can evidently dismiss the alleged observational status of the intuitions that figure in your RE if I can show that they are generated by some explicit theory you hold.

### **Ordinary Belief**

We all collect a lot of beliefs as we go through life being socialized and educated, conversing with friends and strangers, reading books and articles, watching TV, and so on. Many of these beliefs, perhaps even most of them, are true, particularly if we have had a pretty good education and have learned to tell the wheat from the chaff. I am inclined to think a lot of our so-called intuitions about fairness are beliefs of exactly this kind. In the current jargon, they are a part of the "values" we picked up in our families, schools and play groups, our reading and TV watching. It would take a really mad kind of chauvinist, however, to suppose that these have the sort of special epistemological status required to ground a philosophical theory of distributive justice. If I can show that your reflective equilibrium depends on intuitions with this source, I can dismiss it out of hand.

### **Language**

The idea here is that philosophical intuitions are essentially linguistic intuitions about whether, for example, something should be called *fair* or a belief *about H<sub>2</sub>O*. One might hold that they are semantic intuitions on a par with grammatical intuitions, judgments more or less on all fours with judgments about which version of the following sentences are correct:

We are digging a trench to insure/ensure that the spring rains do not wash out the road.

The professor's contemptuous sneer inferred/implied that the student's answer was naive.

More loosely, one might hold that the intuitions in question are part of, or are generated by, the knowledge base one acquires in learning the language.

I have sometimes heard the following argument: "I am a competent speaker of English. A competent speaker of a language is, among other things, someone who can apply its terms correctly. Since 'fair' is a term of English, it follows that I can apply it correctly." If this argument really worked, it would show that every North American child of six could tell beeches from elms. But a competence to use the language correctly is not, more's the pity, an ability to apply its terms truly. One can be a competent speaker of a language and yet regularly misapply many of its terms, especially what we might call its theoretical terms. Holding a might-makes-right theory of fairness will impugn my morals, not my linguistic competence vis-a-vis the word 'fair'. If a child, asked to use 'fair' in a sentence, says, "It isn't fair for girls to get as much as boys," we should suspect the child's politics, not his language.<sup>5</sup>

To harken back to a point made earlier, we can get 'elm' and 'beech' right by deferring to

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

the experts. But we have already seen that the competence this confers cannot be used to shore up RE. Still, I think there is little question that linguistic knowledge is implicated in the generation of philosophical intuitions, and so it is worth asking what kind of linguistic knowledge might be relevant to philosophical intuition.

An account that originated with Davidson (1967) and is championed by Fodor (1975) holds that knowing a language requires knowing satisfaction conditions for its semantically primitive terms. We might store a mentalese sentence equivalent to the following:

'fair' applies to x iff x is fair

Evidently, however, this provides no help in applying 'fair' unless the mentalese equivalent of 'fair', which appears on the right hand side of this biconditional, is either the output of a detector or a pointer to a theory of fairness. Actually, we are going to need a theory either way, since it is pretty obvious that the only way of building a fairness detector is to incorporate a theory of fairness. That theory might be procedural or declarative, or it might be a point in weight space, but the detection of fairness has got to be theoretically mediated. Moreover, the point generalizes: no property that is the target of philosophical intuition is going to be detectable without the mediation of a theory of that property.<sup>6</sup> Any theory of language understanding, then, that involves mapping natural language words onto mentalese equivalents or paraphrases is going to have to assume that applying the natural language term involves accessing a theory of the property the term designates. That theory might be explicit, as it is when we apply a technical scientific term, or it might be tacit, as orthodox cognitive science supposes is the case when we apply ordinary terms to distal stimuli. I have already dealt with explicit theory as the source of philosophical intuition. I will deal with tacit theory below.

A rather less explicit, hence more popular, account of the linguistic knowledge relevant to philosophical intuition is that acquisition of a language involves acquisition of concepts corresponding to the semantic primitives of the language. This version of the idea that linguistic knowledge or competence underlies philosophical intuition, then, reduces to the idea that it is concepts that underlie philosophical intuition. Let us turn, then, to concepts.

## Concepts

There is not a lot of consensus about what concepts are.<sup>7</sup> I think three general approaches can be distinguished, however, all of which lead us to other ground in the context of this discussion.

*Mental Representations.* If we suppose that the concept of fairness is simply a symbol of mentalese meaning *fairness*, then a concept of fairness could generate philosophical intuitions only by functioning as a pointer to some-thing else: an explicit or tacit theory, and hence this idea leads us into no new territory.

*Recognition or Detection Procedures.* This appears to be the core of empiricist accounts of concepts. Since, as we have already seen, detection of the properties targeted by philosophical intuition is bound to be mediated by some tacit or explicit

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

theory, this approach to concepts leads to the same ground we have already mapped out.

*Theories.* The majority view, I think, is that concepts are theories, either explicit, as in the case of technical scientific or legal concepts, or tacit, as in the case of "ordinary" concepts. As emphasized above, for the purposes of the present discussion, I want to construe this option broadly to cover any account that takes the functions of concepts to be mediated by knowledge about the property (or whatever) the concept is a concept of

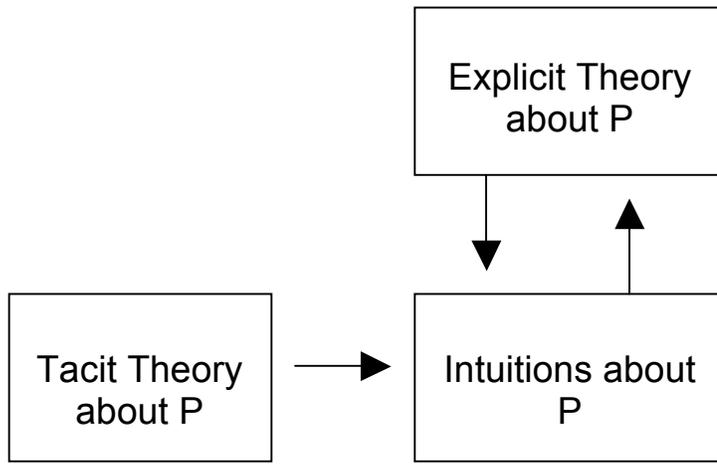
My own view, for what it is worth, is that my concept of an elevator is just everything I know about elevators, different bits of which are activated or accessed on different occasions, depending on cues and previous activations, plus some quick and dirty recognition procedures that account for prototype effects. All that matters for present purposes is that thinking of concepts as the origin of philosophical intuition requires supposing that having the relevant concepts involves having some tacit or explicit theory of the target properties. We can gloss over almost all of the interesting psychological questions here, for these mostly concern what *form* our concepts take. All that matters for my argument is that concepts cannot be what generates philosophical intuitions unless they have a certain *content*: they must be, or provide pointers to, explicit or tacit theories of the target properties.

### **Tacit Theories**

Every remotely plausible story about the origin of philosophical intuitions, then, leads us to the view that they are generated by explicit or tacit theories of the target properties. I have already dealt with the idea that intuitions are generated by explicit theories: There is nothing whatever wrong with judgments of fairness or aboutness being made by the application of an explicit theory. Indeed, I think that is the only respectable way of generating such judgments. But since you cannot suppose that intuitions support the very theories that generate them, this account of the origin of philosophical intuitions undermines RE as a justificatory methodology in philosophy.

Here is the state of play: Philosophical intuitions are generated either by explicit or tacit theory. If the former, they are, perhaps, legitimate, but cannot play the epistemological role required of them by RE. The defender of RE, then, is reduced to the view that philosophical intuition is generated by tacit theory. The picture, then, is this:

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999



It takes no great insight to see that the intuitions, and hence the explicit theories based on them, are only as good as the tacit theory that generates those intuitions. So the viability of RE reduces to the question of the reliability of tacit theories generally, and, in particular, to the question of the reliability of tacit theories of the properties targeted by philosophical intuition. Since the prospects for such theories are very poor, I think we must conclude that RE is not a viable methodology.

What reasons might we have for trusting tacit theories? Descartes argued, in effect, that they are innate, and that what is innate, being from God, cannot be mistaken, since God is no deceiver. I do not think Cartesian theists are a large enough subset of the defenders of RE to worry about. I suspect, however, that some contemporary theorists are tempted to replace God by evolution, arguing that innate theories must be adaptive, since they have survived. I find myself more than a little tempted by the hypothesis that various moral concepts are innate, just as I am tempted by the hypothesis that various mathematical and physical concepts are innate. But Descartes was wrong about innateness being a guarantee of truth. Even the most panglossian reading of evolution will yield at most that innate theories are adaptive, not that they are true. You would have to be either extremely cynical about morals, or extremely naive about social interactions among primates and other social species, to suppose that adaptive moral concepts are accurate. More generally, adaptive theories are *effective* theories, and the

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

effectiveness of a theory is not a simple function of its accuracy. It is also a function of its tractability, and of the kinds of problems it is typically called upon to solve.

We are finite creatures with a fixed cognitive design. Whether a theory is good for us depends a good deal on our design and on the problems we face. How important it is to avoid errors, and which errors to avoid, depends on task and environment. False negatives in food recognition are of little concern when food is plentiful, and they may be a fair price to pay to avoid false positives when there are also lots of poisons around. Conversely false positives in predator recognition are a good price to pay to achieve speed and avoid false negatives. Since accuracy is almost always expensive in terms of time or space or both, cognitive systems often trade accuracy for time and space in cases in which the inaccuracies, though perhaps large, are not serious. Relatively tractable inaccuracy is always better than relatively intractable accuracy provided the inaccuracies are not fatal.

No doubt there is some innate theory. It is pretty likely that we are endowed with some innate physics (Baillargeon, Spelke, and Wasserman 1985; Baillargeon 1987; Spelke 1991). It looks like there might be an innate module responsible for at least some moral reasoning (Cummins 1996a, 1996b). I do not find it at all implausible to suppose that we are endowed with some innate philosophy, especially moral and social philosophy. But when you come to think of what these theories are for, there is no more reason to think that innate philosophy is a good basis for philosophy than that innate physics is a good basis for physics. We have cultural institutions like science and philosophy largely to *overcome* the limitations of our innate endowment. There is, then, no good reason to think that the innate tacit theories that might underlie philosophical intuition are true, and some compelling general reasons for thinking that they are likely to be pretty inaccurate.

Acquired tacit theory fares no better when scrutinized. Undisciplined learning of the sort that produces tacit as opposed to explicit theory is bound, like evolution, to be more guided by effectiveness than by accuracy. Indeed, all learning is guided by effectiveness. It is just that, in explicit educational settings, we take special care that inaccurate theories will not be effective. That is what examinations are all about. We can hardly expect tacit theory to be molded by any such artificial mechanisms, however. Moreover, tacit theory acquisition is bound to exhibit what I call the MCI effect: the verbally and behaviorally expressed views of family and friends are going to loom large in the outcome. They are, in short, going to be biased.

Not only are tacit theories, whether acquired or innate, likely to be inaccurate, they are likely to generate inconsistent intuitions over time. This is partly because one's tacit knowledge can be expected to change over time, but also because which bits of one's tacit knowledge are in fact accessible will vary from occasion to occasion, depending on cues and prior activations. As a consequence, we should expect both order effects and context effects in the intuitions they generate. Perhaps a carefully formulated organon for RE could compensate for such effects, but no one is likely to bother. No one is likely to bother, because, once it is clear what tacit theories are, we are bound to see the intuitions they generate as very likely to be artifacts or errors. And since, as argued in the previous section, we cannot calibrate philosophical intuition unless we are possessed of the sort of theory that would render it idle, there is never going to be any point to compensating for order or context effects in the elicitation of philosophical intuition from tacit theory. If you know enough to start fixing problems with philosophical intuition,

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

you already know enough to get along without it.

## **Capturing the Intuitions**

*Objection:* It is beside the point to argue that philosophical intuitions have no justificatory force, for the point of the enterprise is just to construct a theory that accounts for the intuitions, not to use the intuitions as evidence for that theory.

Of course, there is some interest in knowing what tacit theories we in fact hold, and I suppose something like RE might be a way of at least generating hypotheses about these matters: a theory in RE with intuition might be a reasonable candidate specification of a tacit theory. But, of course, this is simply armchair psychology, and no substitute for empirically disciplined research by people trained to do the job right.

Moreover, achieving RE in some domain is no guarantee that there is a tacit theory at work. The intuitions might be the result of an explicit theory, the results of the RE process itself, seeded, as it were, by some initial tacit or explicit beliefs one happens to have picked up. More seriously, even if there is a tacit theory at work, it is not going to have much psychological interest unless it is pretty universal at some developmental stage or other. Otherwise, it tells us nothing about the architecture of the mind, but simply reflects tacit effects of more or less idiosyncratic experience and indoctrination. Psychology must be careful to distinguish MCI effects from effects of the mind's structure and development. From the point of view of scientific psychology, the fact that some people have internalized a causal theory of reference or a utilitarian theory of distributive justice is of no more interest than the fact that some people have internalized creationism or a Buddhist theory of reincarnation. People get ideas. Unless they are ideas that they always get, irrespective of variations in their environment, psychology may safely leave them to the pages of the *National Enquirer*.

And anyway, the point of RE is not to "capture the intuitions"; the point is to construct and justify a theory of fairness, or aboutness, or some other property of philosophical interest. Knowing what my intuitions about *aboutness* are, and how they are best systematized, is a matter of biographical interest at best. This is often disguised by speaking as if what has been achieved is a systemization of "our" intuitions, the implication being that everyone has them. But, in academic philosophy, at least, this is mainly an illusion created by only allowing fellow travelers to participate in the game. More importantly, even if it were true that everyone shared the intuitions in question, this would have, as we have seen, no tendency to show that those intuitions are accurate.

## **Coherence**

There is, of course, a way of promoting any belief you happen to have into evidence, and

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

that is to embrace some naive form of coherentism. (See Tersman 1993, for just such a defense of RE in ethics.) But just as a plausible coherentism must account for the special role of observation in science, so it must also reveal the disanalogies I have been pressing between intuition and observation. If coherentism cannot tell the epistemological difference between philosophical in-tuition generators and telescopes, then it will fail as an epistemology. For the point of epistemology is to help us to understand such things as the importance of calibration. Trying to save philosophical intuition by adopting coherentism is putting the cart before the horse.

Of course, no coherentism worth its salt requires one to accept one's intuitions, any more than it requires one to accept any other beliefs one happens to have. Hume said that belief is involuntary, that one cannot, for the most part, simply choose what to believe (Hume 1777). You cannot simply decide to believe that you have seven fingers on your right hand. It is a feat completely beyond your powers. I am not sure that all belief is like this, but some certainly is. As far as I can see, there are only two ways a coherence theory can respond to involuntary beliefs. One is to simply allow them their share in determining over-all coherence. The other is to introduce some mechanism for "bracketing" involuntary beliefs that, as it were, one does not want to accept. The first strategy is unattractive because it makes involuntary beliefs tyrannical. Indeed, they will act like foundational beliefs in foundationalist theories, but without any-thing to ground their special epistemological status beyond their involuntariness. It is not surprising, then, that coherence theorists generally adopt some form of the second strategy (Lehrer 1983, 1991). Once a bracketing mechanism is at hand, however, a coherence theorist is free to dismiss RE and philosophical intuition on the grounds that it is, as I have argued, either unnecessary or uncalibratable, and hence does not cohere well with entrenched canons of scientific inference.

## **Conclusion**

Philosophical intuition is epistemologically useless, since it can be calibrated only when it is not needed. Once we are in a position to identify artifacts and errors in intuition, philosophy no longer has any use for it. Moreover, the most plausible account of the origins of philosophical intuitions is that they derive from tacit theories that are very likely to be inaccurate. There is a sense, then, in which philosophical intuitions can always be "explained away": when a dispute arises, I can always, with some plausibility, suppose your intuitions are the artifacts of bad tacit theory. This is a game everyone can play, and I think we should all play it. We should, that is, dismiss philosophical intuition as epistemologically valueless.

To many, this will appear as an unemployment opportunity. But I think there is some hope for a philosophy without intuition. In my field, which is the theory of mental representation, we can, as I mentioned earlier, stop asking whether 'water' refers to H<sub>2</sub>O on Earth but to XYZ on Twin Earth, and start asking what explanatory role representation plays in the theories that appeal to it. This puts the philosophy of psychology on a par with the philosophy of physics. Philosophers of physics interested in space and time do not consult their intuitions any more, they ask how we must understand space and time if the physical theories that appeal to them are to be true and explanatory. I do not know if

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

something analogous will replace intuition in every branch of philosophy, but *something* had better replace it. It cannot support any conclusion worth drawing.

Let me conclude with something about where I think the burden of proof lies. Philosophers, especially those interested in ethics and mathematics, are inclined to argue as follows: "You are no skeptic about ethics and mathematics, so you have to accept the epistemological value of intuition. What else, after all, could ground moral or mathematical knowledge?" I do not know. For that matter, I do not think we have a satisfactory account of perceptual knowledge, either. But "what else" arguments just do not cut it in epistemology. I have just been at pains to show that philosophical theory cannot be grounded in intuition.<sup>8</sup> If, in a given case, we cannot say what it *should* be grounded in, that is no reason for reinstating intuition. It is rather a reason to think harder about the methodology and questions of one's own discipline. I have given up on Twin Earth in the theory of content. Maybe moral philosophers should give up on Trolleys.

## Notes

1. It might be, of course, that the process of achieving RE generates a stable intuition where before there was none, but then the generated intuition has no power to confer an independent check on the theory.

2. This oversimplifies matters a bit. It is not always observations that are the object of scientific prediction. In a sense, any bit of knowledge, whether general or particular, that is not involved in the generation of the theory will do as predictive target. I will return to this point below when it becomes relevant.

3. Could we have perceptions of, say, fairness or aboutness against which philosophical intuition could be tested? If this were possible, RE would not be so central and pervasive. RE looms large in philosophy precisely because no one thinks theories of justice or theories of reference can be perceptually tested.

4. Long regresses, I suspect, will eventually join big circles as examples of vicious structures that are said to become benign if only they are of sufficient magnitude.

5. The phrase 'Linguistic competence' is a technical term of East Pole (Dennett 1986) linguistics. There, competence is opposed to performance, not to incompetence. The competence-performance distinction in linguistics is not an epistemological distinction, it is an expression of the idea that linguistic behavior is best understood via a strategy of idealization. The idea is just that actual linguistic performance does not accurately mirror the underlying grammar, because actual performance is distorted by factors ranging from resource constraints to motivation. Just as we can understand the behavior of a pendulum as the behavior of an "ideal" pendulum—a pendulum in which period depends only on length and gravity—whose behavior is modified by friction and air resistance, so, it is alleged, we can understand the linguistic behavior as the behavior of an ideal speaker—a speaker whose behavior depends only on the underlying grammar—modified by the effects of resource constraints, variable motivation, and the like. Competence in this sense is not the ability to apply terms truly.

6. This is not saying much. Probably no distal property can be detected without the mediation of a certain amount of contingent background knowledge. You might be able to

From Ramsey, W. and M. DePaul (eds.) *The Role of Intuition in Philosophy*. New York: Rowman & Littlefield, pp. 113-127, 1999

build a pretty good P-detector without incorporating any very substantive knowledge of P if there are a range of reliable indicators of P that can be detected al-ready. For example, as Putnam points out (1975), one way to detect beeches and to distinguish them from elms is to ask an expert. So, if you can detect the relevant experts, and you can get them to say whether the target is a beech, you can detect beeches even though you do not have anything approaching the theory that allows the expert to do the trick. Less trivially, you might detect beeches by leaf shape, yet still know relatively little about beeches, and so be at a loss in the winter. I am going to ignore this complication, because all the subsequent argument requires is that the detection of the sort of property that is the target of philosophical intuition must be mediated by some background knowledge.

7. 'Concept' is ambiguous. Generally, a concept is a psychological structure of some sort, but sometimes concepts are taken to be abstract objects that stand to terms as propositions stand to sentences. In the following discussion, concepts are psychological structures.

8. I have heard the following *to quoque*: "Your arguments against appeal to intuition in philosophy are themselves grounded in intuition." I do not think so; I think they are grounded in psychology and successful scientific practice. But here is *a to quoque* back: If you believe in intuition, and think my premises and logic are intuitive, you should accept my conclusion. If you do that, you have *a reduction* against intuition on your hands.