Philos Stud (2007) 132:99–107
DOI 10.1007/s11098-006-9050-3

ORIGINAL PAPER

Experimental philosophy and philosophical intuition

Ernest Sosa

Received: 6 October 2006 / Accepted: 3 November 2006 /
Published online: 14 December 2006

Abstract The topic is experimental philosophy as a naturalistic movement, and its bearing on the value of intuitions in philosophy. This paper explores first how the movement might bear on philosophy more generally, and how it might amount to something novel and promising. Then it turns to one accomplishment repeatedly claimed for it already: namely, the discrediting of armchair intuitions as used in philosophy.

Keywords Experimental philosophy · Intuitions · Casuistry · Normativity · Virtue · Competence

Our topic is experimental philosophy as a naturalistic movement, and its bearing on the value of intuitions in philosophy. I explore first how the movement might bear on philosophy more generally, and how it might amount to something novel and promising. Then I turn to one accomplishment repeatedly claimed for it already: namely, the discrediting of armchair intuitions as used in philosophy.[1]

Experimental philosophy bears on traditional philosophy in at least two ways. It puts in question what is or is not believed *intuitively* by people generally. And it challenges the *truth* of beliefs that *are* generally held, ones traditionally important in philosophy. Each challenge is based on certain experimental results.

How might such experimental results bear on philosophical issues? Here's an example. Traditional skepticism relies crucially on the idea that for all we can really tell, life is but a dream. Whether one enjoys waking life or an extended dream, one has the very same stream of consciousness regardless, so how can one possibly tell the difference? This depends on a conception of dreams as something like hallucinations, however, and we might discover that dreams are not quite like that. Perhaps

[1] This paper was first presented in the ''Experimental Philosophy'' symposium at the 2006 Pacific Division meetings of the APA.

E. Sosa (✉)
Philosophy, Brown University, Box 1918, Providence, RI 02912, USA
e-mail: ernest_sosa@brown.edu

to dream is much more like imagining than like hallucinating. If so, how might this bear on the traditional skeptical problematic?

Even if it *is* part of common sense that in dreams we have conscious experiences intrinsically just like those of waking life, an experimentally based approach might show that common sense is just wrong, in a way that bears crucially on a perennial problematic of philosophy, that of radical skepticism.

Mining the sciences is not in itself novel, of course. Philosophers have been doing that for a very long time, with striking results. Just think of how 20th-century physics bears on the philosophy of space and time, or split-brain phenomena on issues of personal identity, to take just two examples. Perhaps the novelty is rather that experimental philosophers do not so much *borrow* from the scientists as that they *become* scientists. This they do by designing and running experiments aimed to throw light on philosophically interesting issues. And if philosophers are ill-equipped to probe the brain in the ways of neuroscientists, it is easy enough to broaden the movement's self-conception to include interdisciplinary work, provided neuroscientists care enough about such issues with philosophical import, as no doubt some already do. Indeed, many experimental philosophers would already define the movement in this interdisciplinary way.

In any case, most of the actual work so far done in experimental philosophy has involved social psychology. Some of the best-known work has involved surveys designed to probe, and to question, people's intuitions on various philosophical issues. So the novelty might involve the methodologically self-conscious pursuit of such an approach. This anyhow is the side of experimental philosophy that I will be discussing. If the movement is to substantiate a claim to novel results of striking interest to philosophy, this work on intuitions, and other work similarly dependent on surveys, would seem to be particularly important.[2]

My defense against experimentalist objections to armchair intuitions is anchored in the fact that verbal disagreement need not be substantive. This defense will be developed presently, but first: How should we conceive of intuitions?

It is often claimed that analytic philosophy appeals to armchair intuitions in the service of ''conceptual analysis.'' But this is deplorably misleading. The use of intuitions in philosophy should not be tied exclusively to conceptual analysis. Consider some main subjects of prominent debate: utilitarian versus deontological theories in ethics, for example, or Rawls's theory of justice in social and political philosophy, or the externalism/internalism debate in epistemology; and many others could be cited to similar effect. These are not controversies about the conceptual analysis of some concept. They seem moreover to be disputes about something more objective than just a description or analysis of our individual or shared concepts of the relevant phenomena. Yet they have been properly conducted in terms of hypothetical examples, and intuitions about these examples. The questions involved are about rightness, or justice, or epistemic justification. Some such questions concern an ethical or epistemic subject matter, and not just our corresponding concepts.

There can be such a subject matter, beyond our concepts of it, moreover, even if rightness, justice and epistemic justification are not *natural kinds*. Nor need they be socially constructed kinds, either. Indeed, we can regard philosophical controversies

[2] Of course, even if just doing interdisciplinary work with scientists is not surprisingly distinctive or novel, it is still a time-honored tradition, which contemporary experimental philosophy might admirably extend.

as objective without ever going into the ontological status of the entities involved, if any. Mostly we can conduct our controversies, for example, just in terms of where the *truth* lies with regard to them, leaving aside questions of objectual ontology.

Prima facie there is a role for intuition in simple arithmetic and geometry, moreover, but *not only* there. Just consider how extensively we rely on intuition. Take, for example, any two sufficiently different shapes that you perceive on a surface, say the shapes of any two words. If they are words in a foreign language, you may not even have a good *recognitional* grasp, a good concept of any of those shapes. Still you may know perfectly well that they are different. And what you know is not just that the actual *tokens* are different: you also know that any word token *so* shaped *would* be differently shaped from any *thus* shaped (as you demonstrate the two shapes in turn). Or take any shape and any color, or any shape and any sound. And so on, and so forth. Why deny ourselves a similar intuitive access to the simple facts involved in our hypothetical philosophical examples? That we enjoy such access would seem to be the default position, absent some specific objection.

I apply virtue epistemology to the specific case of a priori knowledge, and more specifically to foundational a priori knowledge, to *intuitive* justification and knowledge. Traditionally such intuitions have been understood in accordance with two prominent models: (a) the perceptual, eye-of-the-mind model, and (b) the Cartesian introspective model. Each of these models is subject to fatal objections, however, which prepares the way for my proposed competence-based account.

On my proposal, to intuit that p is to be attracted to assent simply through entertaining that representational content. The intuition is *rational* if and only if it derives from a *competence*, and the content is explicitly or implicitly modal (i.e. attributes necessity or possibility). This first approximation is then defended against the two main published lines of attack on intuitions: the calibration objection, and the cultural divergence objection.[3]

One might quite properly wonder why we should restrict ourselves to modal propositions. And there is no very deep reason. It's just that this seems the proper domain for philosophical uses of intuition. True, contingent intuitions might also derive from a competence. For example, there is a ''taking experience at face value'' competence, whose resulting intuitions would be *of the form* ''if things appear thus and so, then they *are* thus and so.'' These I would call ''empirical'' intuitions, however, to be distinguished from the ''rational'' intuitions involved in abstract, a priori, armchair thought of the kind we do in philosophy.

It might be objected that the proposed account is *too externalist*. But two sources of such worry need to be distinguished. One is the *access* worry, the other the *control* worry, and the two are largely independent. These raise large and fascinating issues of internalism versus externalism. Here I can only gesture, inadequately, at my preferred stance.

First, regarding access, we cannot well insist on armchair access to the justifying power of our sources, since their justifying power depends crucially on their reliability, and this is not knowable from the armchair for our competences generally. (This is not to deny that a source's justifying power is boosted, reaching a special level, when we *do* have access to its reliability; or at least that, when the source

[3] I argue for this approach more fully in earlier papers (Beyer & Burri, 2007; DePaul & Ramsey, 1998; Greenough & Lynch, 2006). And I return to it in Sosa (2007).

operates in combination with such awareness, we attain a higher, reflective level for the resulting beliefs.)

Secondly, regarding control, we cannot well insist on *total* control. We must depend on favorable circumstances in all sorts of ways, and these are often relevantly beyond our control. We must depend on a kind of epistemic luck.

If we insist that true knowledge requires *armchair* access to the reliability of our competences or *total* control regardless of our situation, the outcome is extreme skepticism, which I do not regard as a *reductio* exactly, though I do think it limits the interest of the notions of *absolute* knowledge thus induced.

When we rely on intuitions in philosophy, then, in my view we manifest a competence that enables us to get it right on a certain subject matter, by basing our beliefs on the sheer understanding of their contents. How might survey results create a problem for us? Suppose a subgroup clashes with another on some supposed truth, and suppose they all ostensibly affirm as they do based on the sheer understanding of the content affirmed. We then have a prima facie problem. Suppose half of them affirm <p> while half deny it, with everyone basing their respective attitudes on the sheer understanding of the representational content <p>. Obviously, half of them are getting it right, and half wrong. Of those who get it right, now, how plausible can it be that their beliefs constitute or derive from rational intuition, from an attraction to assent that manifests a real competence?

Not that it is logically incoherent to maintain exactly that. But how plausible can it be, absent some theory of error that will explain why so many are going wrong when we are getting it right? Unless we can cite something different in the conditions or in the constitution of the misled, doubt will surely cloud the claim to competence by those who ex hypothesi are getting it right.

If there is a large disagreement in color judgments within a certain population, how can we sustain the claim to competence by those whose excellent color vision guides them systematically to the truth? Presumably we need to explain the error of the others by appeal to some defect in their lighting conditions or in their color vision, something wrong with their rods and cones or the like. Even if we reject the perceptual model of intuition, so long as we still appeal to competence, we need something analogous to the error theory that protects our color vision from the disagreement of the color blind and of those misled by bad light. We need an error theory that attributes the error of those who disagree with us to bad constitution (blindness) or to bad situation (bad light).

That would seem to be so, moreover, regardless of whether the subject matter is fully objective (as, perhaps, with shape perception), or quasi-objective and reaction-dependent (as, perhaps, with color perception, or with socially constructed phenomena).

So there will definitely be a prima facie problem for the appeal to intuitions in philosophy if surveys show that there is extensive enough disagreement on the subject matter supposedly open to intuitive access.

The bearing of these surveys on traditional philosophical issues is questionable, however, because the experimental results really concern in the first instance only people's responses to certain words. But verbal disagreement *need* not reveal any substantive, real disagreement, if ambiguity and context might account for the verbal divergence. If today I say ''Mary went to the bank yesterday'' and tomorrow you say ''Mary did not go to the bank yesterday'' we need not disagree, given ambiguity and contextual variation. The experimentalists have not yet done enough to show that

they have crossed the gaps created by such potential differences in meaning and context, so as to show that supposedly commonsense intuitive belief is really not as widely shared as philosophers have assumed it to be. Nor has it been shown beyond reasonable doubt that there really are philosophically important disagreements rooted in cultural or socio-economic differences (or so I have argued elsewhere in some detail) (see Bishop & Murphy, 2007).

Within the movement itself, one finds a growing recognition that the supposed ''intuitive disagreements'' may be only verbal. Thus, a recent paper by Shaun Nichols and Joseph Ulatowski contains the following proposal:

> Our hypothesis is that 'intentional' exhibits interpretive diversity, i.e., it admits of different interpretations. Part of the population, when given ... [certain] sorts of cases, interpret 'intentional' one way; and part of the population interpret it in another way. On one interpretation both cases are intentional and on the other interpretation, neither is. In linguistics and philosophy of language, there are several ways that a term can admit of different interpretations: the term might be ambiguous, polysemous, or exhibit certain forms of semantic underspecification. We mean for the interpretive diversity hypothesis to be neutral about which form of interpretive diversity holds for 'intentional'.[4]

To the extent that experimental philosophy adopts this way of accounting for diversity of verbal intuitive responses, it will avoid substantive clashes in favor of merely verbal disagreement. But once such disagreements are seen to be verbal, the supposed problem for philosophical intuition evaporates.

The defense of philosophical intuition by appeal to ''merely verbal disagreement'' may be rejected because the implied failures of communication would threaten to make intuition reports useless for joint philosophical theorizing. Although this point is sometimes pressed, I can see no real threat in it. The appeal to divergence of interpretation is a *defensive* move, made against those who claim that there *is* serious disagreement in supposed intuitions. It is only against such a claim of disagreement that we must appeal to verbal divergence. But any such claim need be taken seriously only when adequately backed by evidence. And this is surely a matter to be taken up case by case. Among possible sources of such attention-demanding evidence, two stand out. First, the evidence might be gathered empirically, through surveys. Second, the evidence might be internal to our field, owed to dialectic with fellow philosophers, where we seem to disagree persistently, for example on what to think about various hypothetical cases. One attractive option, once we have reached that stage, having exhausted other options, would be to consider whether we may be ''interpreting'' our terms somewhat differently.

Consider a further case study of how an apparent clash of intuitions can turn out to be only verbal. We turn to a recent paper by Joshua Knobe and Shaun Nichols about the bearing of intuitions on the problematic of free will and determinism (Knobe & Nichols, forthcoming). In their view intuitions relevant to this problematic are heavily influenced by affect. Here is a brief description of the study, its results, and the morals drawn.

[4] ''Intuitions and Individual Differences: the Knobe Effect Revisited,'' available at http://www.rci.rutgers.edu/~stich/Experimental_Philosophy_Seminar/experimental_philosophy_seminar_readings.htm.

First the distinction between a fully determinist universe D and an indeterminist universe I is presented to experimental subjects, 90% of whom report that our own universe is more like I than like D.

Now for the shocking results: When subjects are asked the abstract question whether agents in D are fully morally responsible, 86% say that they are not: no agent can be fully morally responsible for doing what he is fully determined to do. However, when a dastardly deed is attributed with a wealth of detail to a particular agent in D, and those same subjects are asked whether that agent is then fully morally responsible, 72% report that in their view he is!

Knobe and Nichols consider various ways to account for this amazing divergence. In the end, they find it most plausible to think that some performance error is responsible. Affect, they suggest, degrades intellectual performance in general, whether the relevant competence be memory, perception, inference, etc.

Of course, that explanation will leave intuition affected as lightly as are perception, memory, and inference, unless some further relevant difference can be specified.

In any case, there is an alternative explanation that will cast no affect-involving doubt on the intuitions in play. This other possibility came to mind on reading their paper, and was soon confirmed in the article on moral responsibility in the Stanford Encyclopedia of Philosophy, where we are told that at least two different senses of 'moral responsibility' have emerged: the attributability sense, and the accountability sense.

On the attributability view, to say that S is responsible for action A is to say that A is attributable to S as his own doing, and, we are told in the article, as an action that reveals something about S's character.

On the accountability view, to say that S is responsible for action A is to say that S is properly *held* accountable or responsible for A, in such a way that various good (or bad) things may be visited upon S *for* doing A.

So, here again, quite possibly the striking divergence reported above is explicable mainly if not entirely through verbal divergence.

Indeed, we may plausibly go beyond the explanation suggested in the Stanford Encyclopedia, by suggesting that in common parlance ''accountability'' need *not* be tied to manifestation of character. Those attracted to ''agent causation,'' including philosophers such as Thomas Reid and Roderick Chisholm, would not make that linkage. So, there is a notion of attributability–responsibility that is inherently incompatibilist in requiring only that the agent have caused his action, free of antecedent determinants, *free even of determination by his or her character.*

If so, we may then find different ''interpretations'' at work in the verbal disagreement between the affect-affected intuiters (who react to the specific description of the dastardly deed) and the cold theoretical intuiters (who respond to the abstract question of whether any agent can be responsible in D). Of course, it remains to be seen why the one concept is more readily engaged by the affect elicited with the specific case, and the other more readily by cold reasoning about the abstract issue. But pessimism about explaining this would seem premature.

Let us turn next to a further line of experiment-based objection against philosophical intuition, which appears in a recent paper by Stacey Swain, Joshua Alexander, and Jonathan M. Weinberg, as follows:

> We found that intuitions in response to ... [Keith Lehrer's Truetemp Case] vary according to whether, and what, other thought experiments are considered first. Our results show that: (1) willingness to attribute knowledge in the Truetemp Case increases after being presented with a clear case of non-knowledge, and (2) willingness to attribute knowledge in the Truetemp Case decreases after being presented with a clear case of knowledge. We contend that this instability undermines the supposed evidential status of these intuitions. (Swain, Alexander, & Weinberg, in preparation)

Well, maybe, to *some* extent. But surely the effects of priming, framing, and other such contextual factors will affect the epistemic status of intuition in general, only in the sort of way that they affect the epistemic status of perceptual observation in general. One would think that the ways of preserving the epistemic importance of perception in the face of such effects on perceptual judgments would be analogously available for the preservation of the epistemic importance of intuition in the face of such effects on intuitive judgments. The upshot is that we have to be *careful* in how we use intuition, not that intuition is useless. It is of course helpful to be shown how intuition can go astray in unfavorable conditions, just as perception can go similarly astray. But the important question is untouched: Can intuition enjoy relative to philosophy an evidential status analogous to that enjoyed by perception relative to empirical science?

We turn, finally, to a recent line of attack on philosophical intuition, one also in line with the experimental philosophy movement.[5] According to a recent book by Michael Bishop and J.D. Trout, epistemology should look beyond its navel and adopt the more worthy project of developing prescriptions that will have some use in the real world. By contrast, the methods of ''Standard Analytic Epistemology'' (SAE) ''...are suited to the task of providing an account of the considered epistemic judgments of (mostly) well-off Westerners with Ph.D.'s in Philosophy.'' (Bishop & Trout, 2005).

Normative disciplines concerned with prescription and evaluation have a theoretical side and a more applied side. The latter we might call ''casuistry'' *in a broad sense*. We are familiar with the casuistry of advice columnists, priests, parents, therapists, and friends, tailored to specific individual cases, and we also know the more general, policy-oriented casuistry of applied ethics, a large and thriving subdiscipline. Insofar as there is such a thing as applied epistemology, I suppose it is to be found largely, though not exclusively, in the similarly large and active field of critical reasoning.

It may be objected that even if intuition is defensible abstractly as a possible source of normative knowledge, its role in epistemic casuistry will be small by comparison with our knowledge of the relevant scientific facts about our intellectual equipment and its social and physical setting, about its reliability, and about the reliability of various information-gathering methods.

[5] Of course, not every advocate of ''experimental philosophy'' would endorse everything in the loose conglomerate that falls under that flexible title. Furthermore, there is a recent strain of experimental philosophy with a more positive view of intuitions. Proponents of this strain use experimental evidence to reach a better understanding of those intuitions and of their underlying competence(s). Compare, for examples, the following: Knobe, forthcoming; Nahmias, Morris, Nadelhoffer and Turner, forthcoming; Nichols, 2002 (My thanks here to Joshua Knobe).

That may or may not be so. I find it difficult to assess such size of role, especially since the prospects for epistemic casuistry are so unclear, and I mean epistemic casuistry *as a discipline*, with generally applicable rules. Of course, we know a lot about reliable methods, for example about how to determine a huge variety of facts through the use of a corresponding variety of instruments. And we also know how to use library sources, which newspapers to trust, which statistical methods are reliable, et cetera. But there really is no discernible unified discipline there. Such casuistry would encompass all the manuals for all the various instruments and how to read all the various gauges, for one thing. And it would also include the variegated practical lore on how to tell what's what and on what basis: the lore of navigation, jungle guidance, farming tips, and so on and so forth. That is all of course extremely useful, but it is no part of the traditional problematic of epistemology. Nor is there any reason to replace either of epistemic casuistry or traditional epistemology with the other. Each has its own time and place.

Traditional epistemology enjoys the coherence provided by its unified set of central questions concerning the nature, conditions, and extent of knowledge and justification. Some may regard such questions with distaste. But philistinism is not to be feared by a discipline that has attracted unexcelled minds over the course of millennia, and in cultures as diverse as those of Buddhist India and Classical Greece, and many others.

In any case, even if the role of intuition in epistemic casuistry is small, I fail to see an objection here. Our question has been whether intuition can be understood clearly and defended adequately as a source of foundational a priori justification. Once that is accomplished, our task is completed, especially if intuition's role in epistemic casuistry is *indispensable*, no matter how large or small.

Nevertheless, Bishop and Trout press their case against the theoretical side of SAE, as follows:

> As we have ... argued, when it comes to epistemic judgments, the theories of SAE define what we ''*do* do'' not what we ''*must* or *ought* to do.'' They ... merely tell us how we *do* make epistemic judgments (and by ''we,'' we mean the tiny fraction of the world's population who has studied SAE) ....
>
> The proponent of SAE is replacing normative questions about how to evaluate reason and belief with descriptive questions about how proponents of SAE evaluate reason and belief. (Bishop & Trout, 2005, p. 110)

But this misconstrues the way intuition is supposed to function in epistemology and in philosophy more generally, which is by analogy with the way observation is supposed to function in empirical science.

Empirical theories are required to accord well enough with the deliverances of scientific observation. Does empirical inquiry merely tells us how we *do* make empirical observations? (And by 'we' I mean only the tiny fraction of the world's population who has studied empirical science.) Is the proponent of empirical science replacing questions about the tides, the circulation of the blood, the movements of the planets, and so on, with questions about how proponents of empirical science make certain observations?

That implied parody is supposed to bring out the misconstrual that I find in Bishop and Trout. Intuitions are supposed to function like observations. The data for empirical science includes not *just* claims *about* the observations of some few

specialists. The set of empirical data includes also claims about the subject matter of the specialists' fields of study, about truths concerning the natural phenomena under study. Similarly, philosophical data would include not *just* claims *about* the intuitions shared by some few specialists. Also prominently included would be claims about the subject matter of the philosophers' fields of study, including evaluative or normative truths of epistemology, for example.

Perhaps there is some crucial difference between natural phenomena and evaluative phenomena that rules out any such analogy. Perhaps there are no normative truths, for example, by contrast with the evident availability of empirical truths. But if this is the real issue, then we need to consider whether in principle there could or could not be the truths that there seem intuitively, commonsensically, to be. And how could we possibly approach such a question except philosophically, through the sort of reflection plus dialectic that depends crucially on philosophical intuition?

Even if it turns out that there *is* such a fundamental semantic divergence between empirical and normative subject matter, finally, a relevant analogy between observation and intuition might *still* survive such semantic divergence. This too would need to be debated philosophically. Progress on such issues of metaphilosophy depends thus on progress *within* philosophy.

References

Beyer, C., & Burri, A. (2007). Intuitions: Their nature and epistemic efficacy. *Grazer Philosophische Studien*. *Philosophical Knowledge - Its Possibility and Scope* [Special issue].

Bishop, M., & Murphy, D. (Eds.) (2007). A defense of intuitions. In *Stich and his critics*. Oxford and Boston: Blackwell Publishers. Available at http://homepage.mac.com/ernestsosa/Menu2.html.

Bishop, M. A., & Trout, J. D. (2005). *Epistemology and the psychology of human judgment* (p. 107). Oxford and New York: Oxford University Press.

DePaul, M., & Ramsey, W. (Eds.) (1998). Minimal intuition. In *Rethinking intuition*. New Jersey: Rowman&Littlefield.

Greenough, P., & Lynch, M. (Eds.) (2006). Intuitions and truth. In *Truth and realism.* Oxford and New York: Oxford University Press.

Knobe, J. (forthcoming). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies*.

Knobe, J., & Nichols, S. (forthcoming). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous*.

Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (forthcoming). Is incompatibilism intuitive? *Philosophy and Phenomenological Research*.

Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition, 84*, 221–236.

Sosa, E. (2007). *A virtue epistemology: Apt belief and reflective knowledge*, Vol. 1. Oxford and New York: Oxford University Press.

Swain, S., Alexander, J., & Weinberg, J. M. The instability of philosophical intuitions: Running hot and cold on Truetemp. (in preparation) Available at http://www.indiana.edu/~eel/.