

Daniel C. Dennett

Who's On First?

Heterophenomenology Explained

There is a pattern of miscommunication bedeviling the people working on consciousness that is reminiscent of the classic Abbott and Costello 'Who's on First?' routine. With the best of intentions, people are talking past each other, seeing major disagreements when there are only terminological or tactical preferences — or even just matters of emphasis — that divide the sides. Since some substantive differences also lurk in this confusion, it is well worth trying to sort out. Much of the problem seems to have been caused by some misdirection in my apologia for *heterophenomenology* (Dennett, 1982; 1991), advertised as an explicitly *third*-person approach to human consciousness, so I will try to make amends by first removing those misleading signposts and sending us back to the real issues.

On the face of it, the study of human consciousness involves phenomena that seem to occupy something rather like another dimension: the private, subjective, '*first*-person' dimension. Everybody agrees that this is where we start. What, then, is the relation between the standard 'third-person' objective methodologies for studying meteors or magnets (or human metabolism or bone density), and the methodologies for studying human consciousness? Can the standard methods be extended in such a way as to do justice to the phenomena of human consciousness? Or do we have to find some quite radical or revolutionary alternative science? I have defended the hypothesis that there is a straightforward, conservative extension of objective science that handsomely covers the ground — *all* the ground — of human consciousness, doing justice to all the data without ever having to abandon the rules and constraints of the experimental method that have worked so well in the rest of science. This third-person methodology, dubbed heterophenomenology (phenomenology *of another* not oneself), is, I have claimed, the sound way to take the *first* person point of view as seriously as it can be taken.

Correspondence: D.C. Dennett, Center for Cognitive Studies, Tufts University, Medford, MA 02155-5555, USA. Email: daniel.dennett@tufts.edu

Journal of Consciousness Studies, **10**, No. ??, 2003, pp. ??–??

To place heterophenomenology in context, consider the following ascending scale of methods of scientific investigation:

- experiments conducted on anaesthetized animals;
- experiments conducted on awake animals;
- experiments on human subjects conducted in ‘behaviorese’
 - subjects are treated as much as possible like laboratory rats, trained to criterion with the use of small rewards, with minimal briefing and debriefing, etc.;
- experiments in which human subjects collaborate with experimenters
 - making suggestions, interacting verbally, telling what it is like.

Only the last of these methods holds out much hope of taking human subjectivity seriously, and at first blush it may seem to be a first-person (or, with its emphasis on communicative interaction with the subjects, second-person) methodology, but in fact it is *still* a third-person methodology if conducted properly. It is heterophenomenology.

Most of the method is so obvious and uncontroversial that some scientists are baffled that I would even call it a method: basically, you have to take the vocal sounds emanating from the subjects’ mouths (and your own mouth) and *interpret* them! Well of course. What else could you do? Those sounds aren’t just belches and moans; they’re speech acts, reporting, questioning, correcting, requesting, and so forth. Using such standard speech acts, other events such as button-presses can be set up to be interpreted as speech acts as well, with highly specific meanings and fine temporal resolution. What this interpersonal communication enables you, the investigator, to do is to compose a catalogue of *what the subject believes to be true about his or her conscious experience*. This catalogue of beliefs fleshes out the subject’s *heterophenomenological world*, the world according to S — the subjective world of one subject — not to be confused with the real world. The total set of details of heterophenomenology, plus all the data we can gather about concurrent events in the brains of subjects and in the surrounding environment, comprise the total data set for a theory of human consciousness. It leaves out no objective phenomena and no subjective phenomena of consciousness.

Just what kinds of things does this methodology commit us to? Beyond the unproblematic things all of science is committed to (neurons and electrons, clocks and microscopes, . . .) just to *beliefs* — the beliefs expressed by subjects and deemed constitutive of their subjectivity. And what kind of things are beliefs? Are they sentences in the head written in brain writing? Are they non-physical states of dualist ectoplasm? Are they structures composed of proteins or neural assemblies or electrical fields? We may stay maximally noncommittal about this by adopting, at least for the time being (I recommend: for ever), the position I have defended (Dennett, 1971; 1987; 1991) that treats beliefs from *the intentional stance* as *theorists’ fictions* similar to centres of mass, the Equator, and parallelograms of forces. In short, we may treat beliefs as *abstractions* that measure or describe the complex cognitive state of a subject rather the way

horsepower indirectly but accurately measures the power of engines (don't look in the engine for the horses). As Churchland (1979) has pointed out, physics already has hundreds of well-understood measure predicates, such as *x has weight-in-grams n*, or *x is moving up at n meters per second*, which describe a physical property of *x* by relating it to a *number*. Statements that attribute beliefs using the standard *propositional attitude* format, *x believes that p*, describe *x*'s internal state by relating it to a *proposition*, another kind of useful abstraction, systematized in logic, not arithmetic. We need beliefs anyway for the rest of social science, which is almost entirely conducted in terms of the intentional stance, so this is a conservative exploitation of already quite well-behaved and well-understood methods.

A catalogue of beliefs about experience is not the same as a catalogue of experiences themselves, and it has been objected (Levine, 1994) that 'conscious experiences themselves, not merely our verbal judgments about them, are the primary data to which a theory must answer'. But how, in advance of theory, could we catalogue the experiences themselves? We can see the problem most clearly in terms of a nesting of proximal sources that are presupposed as we work our way up from raw data to heterophenomenological worlds:

- (a) 'conscious experiences themselves'
- (b) beliefs about these experiences
- (c) 'verbal judgments' expressing those beliefs
- (d) utterances of one sort or another

What are the 'primary data'? For heterophenomenologists, the *primary* data are the utterances, the *raw*, uninterpreted data. But before we get to theory, we can interpret these data, carrying us via (c) speech acts to (b) beliefs about experiences.¹ These are the primary *interpreted* data, the pretheoretical data, the *quod erat explicatum* (as organized into heterophenomenological worlds), for a science of consciousness. In the quest for primary data, Levine wants to go all the way to (a) conscious experiences themselves, instead of stopping with (b) subjects' beliefs about their experiences, but this is not a good idea. If (a) outruns (b) — if you have conscious experiences you don't believe you have — those extra conscious experiences are just as inaccessible *to you* as to the external observers. So Levine's proposed alternative garners you no more usable data than heterophenomenology does. Moreover, if (b) outruns (a) — if you believe you have conscious experiences that you don't in fact have — then it is your beliefs that we need to explain, not the non-existent experiences! Sticking to the heterophenomenological standard, then, and treating (b) as the maximal set of primary data, is the way to avoid any commitment to spurious data.

[1] Doesn't interpretation require theory? Only in the minimal sense of presupposing that the entity interpreted is an intentional system, capable of meaningful communication. The task of unifying the interpretation of all the verbal judgments into a heterophenomenological world is akin to reading a novel, in contrast to reading what purports to be true history or biography. The issue of truth and evidence does not arise, and hence the interpretation is as neutral as possible between different theories of what is actually happening in the subject.

But what if some of your beliefs are inexpressible in verbal judgments? If you believe *that*, you can tell us, and we can add that belief to the list of beliefs in our primary data: 'S claims that he has ineffable beliefs about X'. If this belief is true, then we encounter the obligation to explain what these beliefs are and why they are ineffable. If this belief is false, we still have to explain why S believes (falsely) that there are these particular ineffable beliefs. As I put it in *Consciousness Explained*,

You are *not* authoritative about what is happening in you, but only about what *seems* to be happening in you, and we are giving you total, dictatorial authority over the account of how it seems to you, about *what it is like to be you*. And if you complain that some parts of how it seems to you are ineffable, we heterophenomenologists will grant that too. What better grounds could we have for believing that you are unable to describe something than that (1) you don't describe it, and (2) confess that you cannot? Of course you might be lying, but we'll give you the benefit of the doubt (Dennett, 1991, pp. 96–7).

This is all quite obvious, but it has some under-appreciated implications. Exploiting linguistic communication in this way, you get a fine window into the subject's subjectivity but at the cost of a peculiar lapse in normal interpersonal relations. You *reserve judgment* about whether the subject's beliefs, as expressed in their communication, are true, or even well-grounded, but then you treat them as *constitutive* of that subject's subjectivity. (As far as I can see, this is the third-person parallel to Husserl's notion of bracketing or *epoché*, in which the normal presuppositions and inferences of one's own subjective experience are put on hold, as best one can manage, in order to get at the core experience, as theory-neutral and unencumbered as possible.) This interpersonal reserve can be somewhat creepy. To put it fancifully, suppose you burst into my heterophenomenology lab to warn me that the building is on fire. I don't leap to my feet and head for the door; I write down 'subject S believes the building is on fire'. 'No, really, it's on fire!' you insist, and I ask 'Would you like to expand on that? *What is it like* for you to think the building is on fire?' and so forth. In one way I am taking you as seriously as you could ever hope to be taken, but in another way I am not. I am not *assuming* that you are right in what you tell me, but just that that is what you do believe. Of course most of the data-gathering is not done by any such simple interview. Experiments are run in which subjects are prepared by various conversations, hooked up to all manner of apparatus, etc., and carefully debriefed. In short, heterophenomenology is nothing new; it is nothing other than the method that has been used by psychophysicists, cognitive psychologists, clinical neuropsychologists, and just about everybody who has ever purported to study human consciousness in a serious, scientific way.

This point has sometimes been misunderstood by scientists who suppose, quite reasonably, that since I am philosopher I must want to scold somebody for something, and hence must be proposing restrictions on standard scientific method, or discovering limitations therein. On the contrary, I am urging that the prevailing methodology of scientific investigation on human consciousness is not only sound, but readily extendable in non-revolutionary ways to incorporate

all the purported exotica and hard cases of human subjectivity. I want to put the burden of proof on those who insist that third-person science is incapable of grasping the nettle of consciousness.

Let me try to secure the boundaries of the heterophenomenological method more clearly, then, since this has apparently been a cause of confusion. As Anthony Jack has said to me:

It strikes me that heterophenomenology is a method in the same way that ‘empiricism’ is a method, but no more specific nor clearly defined than that. Given how general you seem to allow your definition of heterophenomenology to be, it is no surprise that everything conforms! Perhaps it would be clearer if you explained more clearly what it is supposed to be a counterpoint to — what it is that you object to. I know I am not the only one who has a feeling that you make the goalposts surprisingly wide. So what exactly is a foul? (Jack, personal correspondence).

Lone-wolf autophenomenology, in which the subject and experimenter are one and the same person, is a foul, not because you can’t do it, but because it isn’t science until you turn your self-administered pilot studies into heterophenomenological experiments. It has always been good practice for scientists to put themselves in their own experimental apparatus as informal subjects, to confirm their hunches about what it feels like, and to check for any overlooked or underestimated features of the circumstances that could interfere with their interpretations of their experiments. But scientists have always recognized the need to confirm the insights they have gained from introspection by conducting properly controlled experiments with naive subjects. As long as this obligation is met, whatever insights one may garner from ‘first-person’ investigations fall happily into place in ‘third-person’ heterophenomenology. Purported discoveries that cannot meet this obligation may inspire, guide, motivate, illuminate one’s scientific theory, but *they* are not data — the beliefs of subjects about them are the data. Thus if some phenomenologist becomes convinced by her own (first-)personal experience, however encountered, transformed, reflected upon, of the existence of a feature of consciousness in need of explanation and accommodation within her theory, her *conviction that this is so* is itself a fine datum in need of explanation, by her or by others, but the truth of her conviction must not be presupposed by science.

Does anybody working on consciousness disagree with this? Does anybody think that one can take personal introspection by the investigator as constituting stand-alone evidence (publishable in a peer-reviewed journal, etc.) for any substantive scientific claim about consciousness? I don’t think so. It is taken for granted, so far as I can see, by all the authors in this volume that there is no defensible ‘first-person science’ lying in this quarter, even though that would be the most obvious meaning of the phrase ‘taking a first-person approach’. Thus Cytowic, and Hubbard and Jack, discuss the difficulties in confirming that synaesthesia is more or less what synaesthetes say it is, and never question the requirement that ‘taking the phenomenological reports of these subjects seriously’ (Hubbard and Jack, abstract) requires ‘the personal interaction between subject and experimenter’. And when Hurlburt and Heavey say (abstract) ‘For

example, first-person investigators often rely on questions such as “What were you thinking when you . . . ?” or “How were you feeling when you . . . ?” it apparently does not occur to them that these *aren't* first-person investigations; they are third-person investigations of the special kind that exploit the subject's capacity for verbal communication. They are heterophenomenological inquiries. So I think we can set aside lone-wolf autophenomenology in all its guises. It is not an attractive option, for familiar reasons. The experimenter/subject duality is not what is being challenged by those who want to go beyond the ‘third-person’ methodology. What other alternatives should we consider?

Several critics have supposed that heterophenomenology, as I have described it, is too agnostic or too neutral. Goldman (1997) says that heterophenomenology is not, as I claim, the standard method of consciousness research, since researchers ‘rely substantially on subjects’ introspective beliefs about their conscious experience (or lack thereof)’ (p. 532). In personal correspondence (Feb 21, 2001, available as part of my debate with Chalmers, on my website, at <http://ase.tufts.edu/cogstud/papers/chalmersdeb3dft.htm>) he puts the point this way:

The objection lodged in my paper [Goldman, 1997] to heterophenomenology is that what cognitive scientists *actually* do in this territory is not to practice agnosticism. Instead, they rely substantially on subjects’ introspective beliefs (or reports). So my claim is that the heterophenomenological method is not an accurate description of what cognitive scientists (of consciousness) standardly do. Of course, you can say (and perhaps intended to say, but if so it wasn't entirely clear) that this is what scientists *should* do, not what they *do* do.

I certainly would play the role of reformer if it were necessary, but Goldman is simply mistaken; the adoption of agnosticism is so firmly built into practice these days that it goes without saying, which is perhaps why he missed it. Consider, for instance, the decades-long controversy about mental imagery, starring Roger Shepard, Steven Kosslyn, and Zenon Pylyshyn among many others. It was initiated by the brilliant experiments by Shepard and his students in which subjects were shown pairs of line drawings like the pair in figure 1, and asked to press one

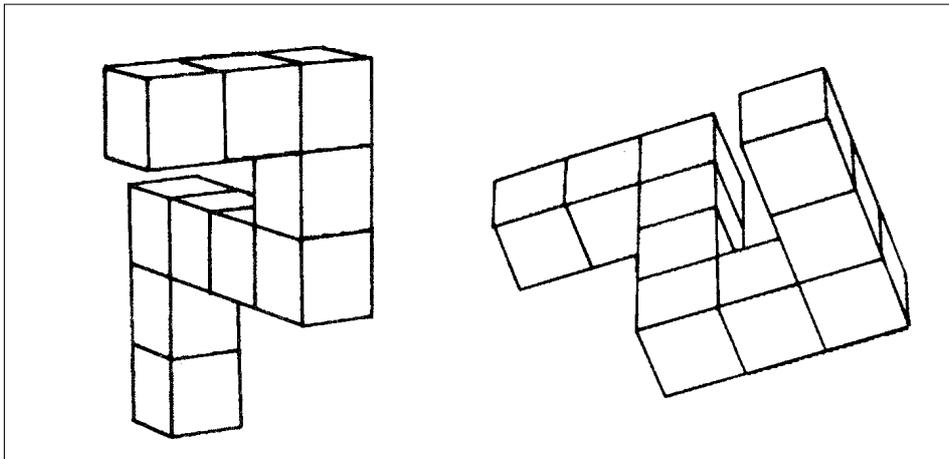


Figure 1

button if the figures were different views of the same object (rotated in space) and another button if they were of different objects. Most subjects claim to solve the problem by rotating one of the two figures in their ‘mind’s eye’ or imagination, to see if it could be superimposed on the other. Were subjects really doing this ‘mental rotation’? By varying the angular distance actually required to rotate the two figures into congruence, and timing the responses, Shepard was able to establish a remarkably regular linear relation between latency of response and angular displacement. Practiced subjects, he reported, are able to rotate such mental images at an angular velocity of roughly 60° per second (Shepard and Metzler, 1971). This didn’t settle the issue, since Pylyshyn and others were quick to compose alternative hypotheses that could account for this striking temporal relationship. Further studies were called for and executed, and the controversy continues to generate new experiments and analysis today (see Pylyshyn, forthcoming, for an excellent survey of the history of this debate; also my commentary, Dennett, forthcoming, both in *Behavioral and Brain Sciences*). Subjects always *say* that they are rotating their mental images, so if agnosticism were not the tacit order of the day, Shepard and Kosslyn would have never needed to do their experiments to support subjects’ claims that what they were doing (at least if described metaphorically) really was a process of image manipulation. Agnosticism is built into all good psychological research with human subjects. In psychophysics, for instance, the use of signal detection theory has been part of the canon since the 1960s, and it specifically commands researchers to control for the fact that the response criterion is under the subject’s control although the subject is not himself or herself a reliable source on the topic. Or consider the voluminous research literature on illusions, both perceptual and cognitive, which standardly assumes that the data are what subjects judge to be the case, and never makes the mistake of ‘relying substantially on subjects’ introspective beliefs’.

The diagnosis of Goldman’s error is particularly clear here: of course experimenters on illusions rely on subjects’ introspective beliefs (as expressed in their judgments) about how it *seems* to them, but that *is* the agnosticism of heterophenomenology; to go beyond it would be, for instance, to assume that in size illusions there really were visual images of different sizes somewhere in subjects’ brains (or minds), which of course no researcher would dream of doing.²

David Chalmers has recently made a similar, if vaguer, claim:

Dennett . . . says scientists have to take a neutral attitude (taking reports themselves as data, but making no claims about their truth), because reports can go wrong. But this misses the natural intermediate option that Max Velmans has called critical phenomenology: accept verbal reports as a *prima facie* guide to a subject’s conscious experience, except where there are specific reasons to doubt their reliability. This seems to be most scientists’ attitude toward verbal reports and consciousness: it’s not ‘uncritical acceptance’, but it’s also far from the ‘neutrality’ of heterophenomenology (Chalmers, 2003).

[2] Goldman has responded to this paragraph in a series of emails to me, which I have included in an Appendix on the website mentioned above.

Chalmers neglects to say how Velmans' critical phenomenology is 'far from' the neutrality of heterophenomenology. I conducted a lengthy correspondence with Velmans on this score and was unable to discover what the purported difference is, beyond Velmans' insisting that his method 'accepts the reality of first-person experience', but since it is unclear what this means, this is something a good scientific method should be agnostic about. Neither Chalmers nor Velmans has responded to my challenge to describe an experiment that is licensed by, or motivated by, or approved by 'critical phenomenology' but off-limits to heterophenomenology, so if there is a difference here, it is one of style or emphasis, not substance. Chalmers has acknowledged this, in a way:

Dennett 'challenges' me to name an experiment that 'transcends' the heterophenomenological method. But of course both views can accommodate experiments equally: every time I say we're using a verbal report or introspective judgment as a guide to first-person data, he can say we're using it as third-person data, and vice versa. So the difference between the views doesn't lie in the range of experiments 'compatible' with them. Rather, it lies in the way that experimental results are interpreted. And I think the interpretation I'm giving (on which reports are given prima facie credence as a guide to conscious experience) is by far the most common attitude among scientists in the field. Witness the debate about unconscious perception among cognitive psychologists about precisely which third-person measures (direct report, discrimination, etc.) are the best guide to the presence of conscious perception. Here, third-person data are being used as a (fallible) guide to first-person data about consciousness, which are of primary interest. On the heterophenomenological view, this debate is without much content: some states subserve report, some subserve discrimination, etc., and that's about all there is to say. I think something like this is Dennett's attitude to those debates, but it's not the attitude of most of the scientists working in the field (Chalmers, 2003).

Chalmers misconstrues my view, as we can see if we look more closely at a particular debate about unconscious perception, to see how heterophenomenology sorts out the issues. Consider *masked priming*. It has been demonstrated in hundreds of different experiments that if you present subjects with a 'priming' stimulus, such as a word or picture flashed briefly on a screen in front of the subject, followed very swiftly by a 'mask' — a blank or sometimes randomly patterned rectangle — before presenting the subjects with a 'target' stimulus to identify or otherwise respond to, there are conditions under which subjects will manifest behaviour that shows they have discriminated the priming stimulus, while they candidly and sincerely report that they were entirely unaware of any such stimulus. For instance, asked to complete the word stem *fri* ____, subjects who have been shown the priming stimulus *cold* are more likely to comply with *frigid* and subjects who have been shown the priming stimulus *scared* are more likely to comply with *fright* or *frightened*, even though both groups of subjects claim not to have seen anything but first a blank rectangle followed by the target to be completed. Now are subjects to be trusted when they say that they were not conscious of the priming stimulus? There are apparently two ways theory can go here:

- A. Subjects are conscious of the priming stimulus and then the mask makes them immediately forget this conscious experience, but it nevertheless influences their later performance on the target.
- B. Subjects unconsciously extract information from the priming stimulus, which is prevented from ‘reaching consciousness’ by the mask.

Chalmers suggests that it is my ‘attitude’ that there is nothing to choose between these two hypotheses, but my point is different. It is open for scientific investigation to develop reasons for preferring one of these theoretical paths to the other, but *at the outset*, heterophenomenology is neutral, leaving the subject’s heterophenomenological worlds bereft of any priming stimuli — that is how it seems to the subjects, after all — while postponing an answer to the question of how or why it seems thus to the subjects. Heterophenomenology is the beginning of a science of consciousness, not the end. It is the organization of the data, a catalogue of *what must be explained*, not itself an explanation or a theory. (This was the original meaning of ‘phenomenology’: a pretheoretical catalogue of the phenomena theory must account for.) And in maintaining this neutrality, it is actually doing justice to the *first-person* perspective, because you yourself, as a subject in a masked priming experiment, cannot discover anything in your experience that favours A or B. (If you think you can discover something — if you notice some glimmer of a hint in the experience, speak up! You’re the subject, and you’re supposed to tell it like it is. Don’t mislead the experimenters by concealing something you discover in your experience. Maybe they’ve set the timing wrong for you. Let them know. But if they’ve done the experiment right, and you really find, so far as you can tell from your own first-person perspective, that you were not conscious of any priming stimulus, then say so, and note that both A and B are still options between which you are powerless to offer any further evidence.)

But now suppose scientists look for a good reason to favour A or B and find it. What could it be? A theory that could provide a good reason would be one that is well-confirmed in other domains or contexts and that distinguishes, say, the *sorts* of discriminations that can be made unconsciously from the sorts that require consciousness. If in this case the masked discrimination was of a feature that in all other circumstances could only be discriminated by a conscious subject, this would be a (fairly) good reason for supposing that, however it may be with other discriminations, in this case the discrimination was conscious-and-then-forgotten, not unconscious. Notice that if anything at all like this were discovered, and used as a ground for distinguishing A from B, it would be a triumph of *third-person* science, not due to anything that is accessible only to the subject’s introspection. Subjects would *learn for the first time* that they were, or were not, conscious of these stimuli when they were taught the theory. It is the neutrality of heterophenomenology that permits such a question to be left open, pending further development of theory. And of course anyone proposing such a theory would have bootstrapped their way to their own proprietary understanding of what they meant by conscious and unconscious subjects, finding a consilience between our everyday assumptions about what we are conscious of and what we are not, on

the one hand, and their own classificatory scheme on the other. Anything too extreme ('It turns out on our theory that most people are conscious for only a few seconds a day, and nobody is conscious of sounds at all; hearing is entirely unconscious perception') will be rightly dismissed as an abuse of common understanding of the terms, but a theory that is predictively fecund and elegant can motivate substantial abandonment of this anchoring lore. Only when such a theory is in place will we be able, for the first time, to *know what we mean* when we talk about 'the experiences themselves' as distinct from what we each, subjectively, take our experiences to be.

This sketches a clear path to settling the issue between A and B, or to discovering good reasons for declaring the question ill-posed. If Chalmers thinks that scientists do, and should, prefer a different attitude towards such questions, he should describe in some detail what it is and why it is preferable. In fact, I think that while there has been some confusion on this score (and some spinning of wheels about just what would count as favouring unconscious perception over conscious perception with forgetting), scientists are comfortable with the heterophenomenological standards.

Varela and Shear (1999) describe the empathy of the experimenter that they see as the distinguishing feature of a method they describe as first-person:

In fact, that is how he sees his role: as an empathic resonator with experiences that are familiar to him and which find in himself a resonant chord. This empathic position is still partly heterophenomenological, since a modicum of critical distance and of critical evaluation is necessary, but the intention is entirely other: to meet on the same ground, as members of the same kind. . . . Such encounters would not be possible without the mediator being steeped in the domain of experiences under examination, as nothing can replace that first-hand knowledge. This, then, is a radically different style of validation from the others we have discussed so far (p. 10).

One can hardly quarrel with the recommendation that the experimenter be 'steeped in the domain of experiences' under examination, but is there more to this empathy than just good, knowledgeable interpretation? If so, what is it? In a supporting paper, Thompson speaks of 'sensual empathy', and opines: 'Clearly, for this kind of sensual empathy to be possible, one's own body and the Other's body must be of a similar type' (2001, p. 33). This may be clear to Thompson, but in fact it raises a highly contentious set of questions: Can women not conduct research on the consciousness of men? Can slender investigators not explore the phenomenology of the obese? Perhaps more to the point, can researchers with no musical training or experience ('tin ears') effectively conduct experiments on the phenomenology of musicians? When guidance from experts is available, one should certainly avail oneself of it, but the claim that one must *be* an expert (an expert musician, an expert woman, an expert obese person) before conducting the research is an extravagant one. Suppose, however, that it is true. If so, we should be able to discover this by attempting, and detectably failing, to conduct the research as well as the relevant experts conduct the research. *That* discovery would itself be something that could only be made by first adopting the neutral heterophenomenological method and then assaying the results in comparison

studies. So once again, the neutral course to pursue is not to *assume* that men can't investigate the consciousness of women, etc., but to investigate the question of whether we can discover any good scientific reason to believe this. If we can, then we should adjust the standards of heterophenomenology accordingly. It is just common sense to design one's experiments in such a way as to minimize interference and maximize efficiency and acuity of data-gathering.

Is there, then, any 'radically different style of validation' on offer in these proposals? I cannot find any. Some are uneasy about the noncommittal stance of the heterophenomenologist. Wouldn't the cultivation of deep trust between subject and experimenter be better? Apparently not. The history of *folie à deux* and Clever Hans phenomena suggests that quite unwittingly the experimenter and the subject may reinforce each other into artifactual mutual beliefs that evaporate when properly probed. But we can explore the question. It is certainly wise for the experimenter not to antagonize subjects, and to encourage an atmosphere of 'trust' — note the scare quotes. The question is whether experimenters should go beyond this and *actually trust* their subjects, or should instead (as in standard experimental practice) quietly erect the usual barriers and foils that keep subjects from too intimate an appreciation of what the experimenters have in mind. Trust is a two-way street, surely, and the experimenter who gets in a position where the subject can do the manipulating has lost control of the investigation.

I suspect that some of the dissatisfaction with heterophenomenology that has been expressed is due to my not having elaborated fully enough the potential resources of this methodology. There are surely many subtleties of heterophenomenological method that have yet to be fully canvassed. The policy of training subjects, in spite of its uneven history in the early days of psychology, may yet yield some valuable wrinkles. For instance, it might in some circumstances heighten the powers of subjects to articulate or otherwise manifest their subjectivity to investigators. The use of closed loop procedures, in which subjects to some degree control the timing and other properties of the stimuli they receive is another promising avenue. But these are not alternatives to heterophenomenology, which is, after all, just the conservative extension of standard scientific methods to data gathering from awake, communicating subjects.

Why *not* live by the heterophenomenological rules? It is important to appreciate that the reluctance to acquiesce in heterophenomenology as one's method is ideology-driven, not data-driven. Nobody has yet pointed to any variety of data that are inaccessible to heterophenomenology. Instead, they have objected 'in principle', perhaps playing a little gorgeous Bach for the audience and then asking the rhetorical question, 'Can anybody seriously believe that the wonders of human consciousness can be exhaustively plumbed by *third-person methods*??' Those who are tempted to pose this question should either temper their incredulity for the time being or put their money where their mouth is by providing the scientific world with some phenomena that defy such methods, or by describing some experiments that are clearly worth doing but that would be ruled out by heterophenomenology. I suspect that some of the antagonism to heterophenomenology is generated by the fact that the very neutrality of the methodology

opens the door to a wide spectrum of theories, including some — such as my own — that are surprisingly austere, deflationary theories according to which consciousness is more like stage magic than black magic, requiring no revolution in either physics or metaphysics. Some opponents to heterophenomenology seem intent on building the mystery into the very setting of the problem, so that such deflationary theories are disqualified at the outset. Winning by philosophical footwork what ought to be won by empirical demonstration has, as Bertrand Russell famously remarked, all the advantages of theft over honest toil. A more constructive approach recognizes the neutrality of heterophenomenology and accepts the challenge of demonstrating, empirically, in its terms, that there are marvels of consciousness that cannot be captured by conservative theories.

References

- Chalmers, David J. (2003), 'Responses to articles on my work' <http://www.u.arizona.edu/~chalmers/responses.html#dennett2>
- Churchland, Paul M. (1979), *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press).
- Dennett, Daniel C. (1971), 'Intentional Systems', *J. Phil.*, **68**, pp. 87–106.
- Dennett, Daniel C. (1982), 'How to study consciousness empirically, or Nothing comes to mind', *Synthese*, **59**, pp. 159–80.
- Dennett, Daniel C. (1987), *The Intentional Stance* (Cambridge, MA: MIT Press/Bradford).
- Dennett, Daniel C. (1991), *Consciousness Explained* (Boston, MA: Little Brown).
- Dennett, Daniel C. (forthcoming), 'Does your brain use the images in it, and if so, how?' commentary on Pylyshyn (forthcoming).
- Goldman, Alvin (1997), 'Science, Publicity and Consciousness', *Philosophy of Science*, **64**, pp. 525–45.
- Levine, Joseph (1994), 'Out of the closet: A qualophile confronts qualophobia', *Philosophical Topics*, **22**, pp. 107–26.
- Pylyshyn, Zenon W. (forthcoming), 'Mental Imagery: In search of a theory', Target article in *Behavioral and Brain Sciences*.
- Shepard, R.N. and Metzler, J. (1971), 'Mental rotation of three-dimensional objects', *Science*, **171**, pp. 701–3.
- Thompson, Evan (2001), 'Empathy and consciousness', *Journal of Consciousness Studies*, **8** (5–7), pp. 1–33.
- Varela, Francisco and Shear, Jonathan (1999), 'First-person methodologies: What, Why, How?', *Journal of Consciousness Studies*, **6** (2–3), pp. 1–14.