

## Chapter Seven

### The Unreliability of Naive Introspection

Is it not one and the same “I” who is now doubting almost everything, who nonetheless understands some things, who affirms that this one thing is true, denies everything else, desires to know more, is unwilling to be deceived, imagines many things even involuntarily, and is aware of many things which apparently come from the senses? Are not all these things just as true as the fact that I exist, even if I am asleep all the time, and even if he who created me is doing all he can to deceive me?... Lastly, it is also the same “I” who has sensory perceptions, or is aware of bodily things as it were through the senses. For example, I am now seeing light, hearing a noise, feeling heat. But I am asleep, so all this is false. Yet I certainly *seem* to see, to hear, and to be warmed. This cannot be false....

– René Descartes, *Meditations on First Philosophy* (1641/1984), p. 19

i.

Current conscious experience is generally the last refuge of the skeptic against uncertainty. Though we might doubt the existence of other minds, that the sun will rise tomorrow, that the earth existed five minutes ago, that there’s any “external world” at all, still we can know, it’s said, the basic features of our ongoing stream of experience.

Descartes espouses this view in his first two *Meditations*. So does Hume, in the first book of the *Treatise*, and – as I read him – the ancient skeptic Sextus Empiricus.<sup>1</sup> Other

radical skeptics like Zhuangzi and Montaigne, though they appear to aim at very general skeptical goals, don't grapple specifically and directly with the possibility of radical mistakes about current conscious experience. Is this an unmentioned exception to their skepticism? Unintentional oversight? Do they dodge the issue for fear that it is too poor a field on which to fight their battles?<sup>2</sup> Where is the skeptic who says: We have no reliable means of learning about our own ongoing conscious experience, our current imagery, our inward sensations – we are as in the dark about that as about anything else, perhaps even more in the dark?

Is introspection (if that's what's going on here) just that good? If so, that would be great news for the blossoming – or should I say recently resurrected? – field of consciousness studies. Or does contemporary discord about consciousness – not just about the *physical bases* of consciousness but seemingly about the basic features of experience itself – point to some deeper, maybe fundamental, elusiveness that somehow escaped the notice of the skeptics, that perhaps partly explains the first, ignoble death of consciousness studies a century ago?

ii.

One must go surprisingly far afield to find major thinkers who unambiguously hold, as I do, that the introspection of current conscious experience is both (i.) possible, important, and central to the development of a full scientific understanding of the mind, and (ii.) highly untrustworthy, at least as commonly practiced. In some Eastern meditative traditions, I think this conjunction is a commonplace. Also the fiercest advocates of

introspective training in the first era of scientific psychology (circa 1900) endorsed both claims – especially E.B. Titchener (see Chapter 5; for a brief discussion of the early “phenomenologists” see this note<sup>3</sup>). Both the meditators and Titchener, though, express optimism about introspection “properly” conducted – so they hardly qualify as *general* skeptics or pessimists. It’s as though their advocacy of a regimen set them free to criticize introspection as ordinarily practiced. But might they be right more in their doubts than in their hopes? Might understanding the human mind require introspection, though the prospects are bleak?

I won’t say much to defend (i), which I take to be both common sense and the majority view in philosophy. Of course we have *some* sort of attunement to our ongoing conscious experience, and we impoverish ourselves if we try to do without it. Defending (ii) is my project. In less abbreviated form: Most people are poor introspectors of their own ongoing conscious experience. We fail not just in assessing the *causes* of our mental states or the processes underwriting them; and not just in our judgments about nonphenomenal mental states like traits, motives and skills; and not just in determining the proper labels for or essences of states otherwise perfectly well known; and not only when we are distracted, or passionate, or inattentive, or self-deceived, or pathologically deluded, or when we’re reflecting about minor matters, or about the past, or only for a moment, or where fine discrimination is required. We are both ignorant and prone to error. There are major lacunae in our self-knowledge that are not easily filled in, and we make gross, enduring mistakes about even the most basic features of our currently ongoing conscious experience, even in favorable circumstances of careful reflection, with

distressing regularity. We either err or stand perplexed, depending – rather superficially, I suspect – on our mood and caution.

Contemporary philosophers and psychologists often doubt the layperson's talent in assessing such nonconscious mental states as personality traits, motivations and skills, hidden beliefs and desires, and the bases of decisions; and they may describe such doubts as doubts about "introspection". But it's one thing not to know why you chose a particular pair of socks (to use an example from Nisbett and Wilson 1977) and quite another to be unable accurately to determine your currently ongoing *visual experience* as you look at those socks, your *auditory experience* as the interviewer poses the question, the experience of pain in your back making you want to sit down. Few philosophers or psychologists express plain and general pessimism about the latter sorts of judgment. Or, rather, I should say this: I have heard such pessimism mainly from *behaviorists*, and their near cousins, who nest their arguments in a theoretical perspective that rejects the psychological value, sometimes even the coherence, of attempting to introspect conscious experience at all – and thus reject (i) above – though indeed even radical behaviorists often pull their punches when it comes to ascribing flat out *error* (see Watson 1913; Skinner 1945, 1953; Ryle 1949; Bem 1972<sup>4</sup>).

Accordingly, though infallibilism – the view that we cannot err in our judgments about current conscious experience – is now mostly out of favor, mainstream philosophical criticism of it is meek. Postulated mistakes are largely only momentary, or about matters of fine detail, or under conditions of stress or pathology, or at the hands of malevolent neurosurgeons.<sup>5</sup> Fallibilists generally continue to assume that, in favorable circumstances, careful introspection can reliably reveal at least the broad outlines of

one's currently ongoing experience. Even philosophers most of the community sees as pessimistic about introspection are, by my lights, remarkably tame and generous. Paul Churchland, famous for his disdain of ordinary people's theories about the mind, nevertheless puts the accuracy of introspection on a par with the accuracy of sense perception (1985, 1988). Daniel Dennett, who seems in some places to offer arguments for introspective error, in other places says that we can come close to infallibility when charitably interpreted.<sup>6</sup> Where are the firebrands?

A word about "introspection": I happen to regard it as a species of attention to currently ongoing conscious experience, but I won't defend that view in this book (see Schwitzgebel forthcoming-a). My project here, I think, stands or falls quite independently of that particular account. Or rather: I regard the reflections of this chapter and book as *constraining* any viable theory of introspection. That is, we can and should consider the kinds of mistakes that people make about their stream of experience before we construct a detailed theory of introspection; it is only in light of conclusions about reliability that theories of the mechanisms can properly be developed. Think of introspection as you will, then, as long as it is the primary method by which we normally reach judgments about our experience in cases of the sort I'll describe.<sup>7</sup> That method, whatever it is, is unreliable as typically executed. Or so I'll argue.

iii.

I don't know what emotion is, exactly. Neither do you, I'd guess. Is surprise an emotion? Comfort? Irritability? Is it more of a gut thing or a cognitive thing?

Assuming cognition isn't totally irrelevant, how is it involved? Does cognition relate to emotion merely as cause and effect, or is it somehow, partly, constitutive? I'm not sure there's a single right answer to these questions. The empirical facts seem ambiguous and tangled (for a recent review, see Prinz 2004). Probably we need to conjecture and stipulate, simplify, idealize, to have anything workable. So also, probably, for most interesting psychological concepts. But here's one thing that's clear: Whatever emotion is, some emotions – joy, anger, fear – can involve or accompany conscious experience.

Now you're interested, presumably, in philosophy and psychology, in introspection, consciousness, and the like, or you wouldn't be reading this book. You've had emotional experiences, and you've thought about them, reflected on how they feel as they've been ongoing or in the cooling moments as they fade. If such experiences are introspectible, and if introspection is the diamond clockwork often supposed, then you have some insight. So tell me: Are emotional states like joy, anger, and fear always felt phenomenally – that is, as part of one's stream of conscious experience – or only sometimes? Is their phenomenology, their experiential character, always more or less the same, or does it differ widely from case to case? For example, is joy sometimes in the head, sometimes more visceral, sometimes a thrill, sometimes an expansiveness – or, instead, does joy have a single, consistent core, a distinctive, identifiable, unique experiential character? Is emotional consciousness simply the experience of one's bodily arousal and other bodily states, as William James (1890/1981) and Jesse Prinz (2004) suggest? Or, as most people think, can it include, or even be exhausted by, something less literally visceral? Is emotional experience consistently located in space – for example, in particular places in the interior of one's head and body? Can it have color?

For instance, do we sometimes literally “see red” as part of being angry? Does it typically come and pass in a few moments (as Buddhist meditators sometimes suggest), or does it tend to last awhile (as my English-speaking friends more commonly say)?

If you’re like me, you won’t find all such questions trivially easy. You’ll agree that someone – maybe even you yourself – could be mistaken about some of them, despite sincerely attempting to answer them, despite a history of introspection, despite maybe years of psychotherapy or meditation or self-reflection. You can’t answer these questions one-two-three with the same easy confidence that you can answer similarly basic structural questions about cars – how many wheels? hitched to horses? travel on water? If you can – well, I won’t try to prove you wrong! But if my past inquiries are indicative, you are in a distinct minority.

It’s not just *language* that fails us – most of us? – when we confront such questions (and if it were, we’d have to ask, anyway, why this particular linguistic deficiency?) but introspection itself. The questions challenge us not simply because we struggle for the *words* to describe a patently obvious phenomenology. It’s not like perfectly well knowing what particular shade of tangerine your Volvo is, stumped only about how to describe it. Rather, in the case of emotion, the very phenomenology itself – the “qualitative” character of our consciousness – is not entirely evident, or so it seems to me. But how could this be so, if we know the “inner world” of our own experience so much better than the world outside? Even the grossest features of our emotional experience largely elude us. Reflection doesn’t remove our ignorance, or it delivers haphazard results.

Relatedly, most of us have a pretty poor sense, I suspect, of what brings us pleasure and suffering. Do you really enjoy Christmas? Do you really feel bad while doing the dishes? Are you happier weeding or going to a restaurant with your family? Few people make a serious study of this aspect of their lives, despite the lip service we generally pay to the importance of “happiness”. Most people feel bad a substantial proportion of the time, it seems to me. We are remarkably poor stewards of our emotional experience. We may *say* we’re happy – overwhelmingly we do – but we have little idea what we’re talking about.<sup>8</sup>

iv.

Still, you might suggest, when we attend to *particular instances* of ongoing emotional experience, we can’t go wrong, or don’t, or not by far. We may concede the past to the skeptic, but not the present. It’s impossible – nearly impossible? – to imagine my being wrong about my ongoing conscious experience *right now*, as I diligently reflect.

Well, philosophers sometimes *say* this, or things that sound a lot like it – Descartes, for example, in the epigraph, as standardly interpreted<sup>9</sup> – but I confess to wondering whether they’ve really thought it through, contemplated a variety of examples, challenged themselves. You’d hope they would have, so maybe I’m misunderstanding or going wrong in some way here. But to me at least, on reflection, the claim that I could be infallible, or nearly so, in all or most of what I’m inclined to say about my ongoing consciousness – even barring purely linguistic errors, and even assuming I’m being diligent and cautious and restricting myself to purely phenomenal claims about present



experience, arrived at (as far as I can tell) “introspectively” – well, unfortunately, that seems just blatantly unrealistic.

Let’s try an experiment. You’re the subject. Reflect on, introspect, your own ongoing emotional experience right now. Do you even have any? If you’re in doubt, vividly recall some event that still riles you until you’re sure enough that you’re suffering some renewed emotion. Or maybe your boredom, anxiety, irritation, or whatever in reading this essay is enough. Now let me ask: Is it completely obvious to you what the character of that experience is? Does introspection reveal it to you as clearly as visual observation reveals the presence of the text before your eyes? Can you, through introspection, discern the gross and fine features of your emotional phenomenology as easily and confidently as you can, through vision, discern the gross and fine features of nearby external objects? Can you trace its spatiality (or nonspatiality), its viscerality or cognitiveness, its involvement with conscious imagery, thought, proprioception, or whatever, as sharply and infallibly as you can discern the shape, texture, and color of your desk? (Or the difference between 3 and 27?) I cannot, of course, force a particular answer to these questions. I can only invite you to share my intuitive sense of uncertainty.

Or consider this: My wife mentions that I seem to be angry about being stuck with the dishes again (despite the fact that doing the dishes makes me happy?). I deny it. I reflect; I sincerely attempt to discover whether I’m angry – I don’t just reflexively defend myself but try to be the good self-psychologist my wife would like me to be – and I still don’t see it. I don’t think I’m angry. But I’m wrong, of course, as I usually am in such situations: My wife reads my face better than I introspect. Maybe I’m not quite *boiling*

inside, but there's plenty of angry phenomenology to be discovered if I knew better how to look. Or do you think that every time we're wrong about our emotions, those emotions must be nonconscious, dispositional, not genuinely *felt*? Or felt and perfectly apprehended phenomenologically but somehow nonetheless mislabeled? Can't I also err more directly?

Surely my denial of anger is colored by my wanting to maintain a particular self-conception and by the same involvement in the situation that produces the anger itself. To that extent, it's less than ideal as a test of my claim that, even in the most favorable circumstances of quiet reflection, we are prone to err about our experience. However, as long as we focus on judgments about emotional phenomenology, we can't insist that the reflection always be cool. If that's enough consistently to undermine the reliability of our judgments, that rather better supports my thesis than defeats it, I think.

Excellent, nearly infallible, judges of our emotional phenomenology? I'm baffled. How could anyone believe that? Do *you* believe that?

v.

Now maybe emotional experience is an unusually difficult case. Maybe, though we err there, we are generally quite accurate in our judgments about *other* aspects of our phenomenology. Maybe my argument even plays on some conceptual confusion about the relation between emotion and its phenomenology or relies illegitimately on introspection's undercutting the emotion introspected. I don't think so, but I confess I have no tidy account to eradicate such worries.

So let's try vision. Suppose I'm looking directly at a nearby, bright red object in good light, and I judge that I'm having the visual phenomenology, the "inward experience", of redness. Here, perhaps, even if not in the case of emotion, it seems rather hard to imagine that I could be wrong (though I could be wrong in using the term "red" to *label* an experience I otherwise perfectly well know).

I'll grant that. Some aspects of visual experience are so obvious it would be difficult to go wrong about them. So also would it be difficult to go wrong in some of our judgments about the external world – the presence of the text before your eyes, the existence of the chair in which you're sitting and now (let's suppose) minutely examining. Introspection may admit obvious cases, but that in no way proves that it's more secure than external perception, or even as secure.

Now of course many philosophers have argued plausibly that one *could* be wrong even in "obvious" judgments about external objects, if one allows that one may be dreaming or that one's brain may have been removed at night and teleported to Alpha Centauri to be stimulated by genius neuroscientists with inputs mimicking normal interaction with the world. Generally, philosophers have supposed, with Descartes, that such thought experiments don't undermine judgments about visual phenomenology. Brain-in-a-vat skepticism suggests that I may be wrong about seeing a red tomato, but it doesn't undermine my introspective knowledge that I am having *an experience as of* seeing a red tomato. So perhaps obvious introspective judgments (about "inner" experience) *are* more secure than obvious perceptual ones (about how objects stand in the world outside), since they don't admit even this particular smidgen – usually it only seems like a smidgen – of doubt?

But in dreams we make baldly incoherent judgments, or at least very stupid ones. I think I can protrude my tongue without its coming out; I think I see red carpet that's not red; I see a seal as my sister without noticing any difficulty about that. In dream delirium, these judgments may seem quite ordinary or even insightful. If you admit the possibility that you're dreaming, I think you should admit the possibility that your judgment that you're having red phenomenology is a piece of delirium, unaccompanied by any actual reddish phenomenology. Indeed, it seems to me not entirely preposterous to suppose that we have no color experiences at all in our sleep, or have them only rarely, and our judgments about the colors of dream objects are on par with the seal-sister judgment, purely creative fiction unsupported by any distinctive phenomenology. (See Chapter 1 for more skepticism about color in dreams.) If so, the corresponding judgments about the coloration of our *experiences* of those dream objects will be equally unsupported.

Likewise, if malevolent neurosurgeons from Alpha Centauri might massage and stoke our brains, I see no reason to deny them the power to produce directly the judgment that one is having reddish phenomenology, while suppressing the reddish phenomenology itself. Is this so patently impossible?<sup>10</sup>

*Absolute* security and immunity to skeptical doubt thus elude even "obvious" introspective judgments as well as perceptual ones. If we exclude radically skeptical worries, then we're left with judgments on a par ("red phenomenology now", "book in my hands") – judgments as obvious and secure as one could reasonably wish. The issue of whether the introspection of current visual experience warrants greater trust than the perception of nearby objects must be decided on different grounds.

Look around a bit. Consider your visual experience as you do so. Does it seem to have a center and a periphery, differing somehow in clarity, precision of shape and color, richness of detail? Yes? It seems that way to me too. Now consider this: How *broad* is that field of clarity? Thirty degrees? More? Maybe you're looking at your desk, as I am. Does it seem that a fairly wide swath of the desk – a square foot? – presents itself clearly in experience at any one moment, with the shapes, colors, textures all sharply defined? Most people endorse something like this when I ask them.<sup>11</sup> They are, I think, mistaken.

Consider, first, our visual capacities. It's firmly established that the precision with which we detect shape and color declines precipitously outside a central, foveal area of about one to two degrees of arc (about the size of your thumbnail held at arm's length). Dennett (1991) has suggested a way of demonstrating this to yourself. Draw a card from a normal deck without looking at it. Keeping your eyes fixed on some point in front of you, hold the card to one side, at arm's length, just beyond your field of view. Without moving your eyes, slowly rotate the card toward the center of your visual field. How close to the center must you bring it before you can determine the color of the card, its suit, and its value? Most people are quite surprised at the result of this little experiment. They substantially overestimate their visual acuity outside the central, foveal region. When they can't make out whether it's a Jack, Queen, or King, though the card is nearly (but only *nearly*) dead center, they laugh, they're astounded, dismayed (see also Dennett 2001).

By itself, this says nothing about our visual experience. Surprise and dismay may reveal error in our normal (implicit) assumptions about our visual capacities, but it's one thing to mistake one's abilities and quite another to misconstrue phenomenology (a point also emphasized in Chapter 4's discussion of echolocation). Our visual experience depends on the recent past, on general knowledge, on what we hear, think, and infer, as well as on immediate visual input – or so it's plausible to suppose. Background knowledge could thus fill in and sharpen our experience beyond the narrow foveal center. Holding our eyes still and inducing ignorance might artificially crimp the region of clarity.

Still, I doubt visual experience is nearly as sharp and detailed as most untutored introspectors seem to think. Here's the root of the mistake, I suspect: When the thought occurs to you to reflect on some part of your visual phenomenology, you normally *move your eyes* (or “foveate”) in that direction. Consequently, wherever you think to attend, within a certain range of natural foveal movement, you find the clarity and precision of foveal vision. It's as though you look at your desk and ask yourself: Is the stapler clear? Yes. The pen? The artificial wood grain between them and the mouse pad? Yes – each time looking directly at the object in question – and then you conclude that they're all clear simultaneously. (This, then, would be a version of the “refrigerator light error” discussed in Chapter 6; see also Dennett 1969, p. 139-141. Another contributing factor may be an implicit analogy between visual experience and painting or photography; see Chapter 2 and Noë 2004.)

But you *needn't* reflect in this way. We can pry foveation apart from introspective attention. Fixate on some point in the distance, holding your eyes steady

while you reflect on your visual experience outside the narrow fovea. Better, direct your introspective energies away from the fovea while your eyes continue to move around (or “saccade”) normally. This may require a bit of practice. You could start by keeping one part of your visual field steadily in mind while you allow your eyes to foveate anywhere but there. Take a book in your hands and let your eyes saccade around its cover, while you think about your visual experience in the regions away from the precise points of foveation.

Most of the people I’ve spoken to, who attempt these exercises, eventually conclude to their surprise that their experience of clarity decreases substantially even a few degrees from the center. Through more careful and thoughtful introspection, they seem to discover – I think they really do discover – that visual experience does not consist of a broad, stable field, flush with precise detail, hazy only at the borders. They discover that, instead, the center of clarity is tiny, shifting rapidly around a rather indistinct background. My interlocutors – most of them – confess to error in having originally thought otherwise.<sup>12</sup>

If I’m right about this, then most naive introspectors are badly mistaken about their visual phenomenology when they first reflect on it. Even though they may be patiently considering their experience as it occurs, they will tend to go wrong unless they are warned and coached against a certain sort of error. And the error they make is not a subtle one; the two conceptions of visual experience differ vastly. If naive introspectors are as wrong as many later confess themselves to be, they’re wrong about an *absolutely fundamental and pervasive* aspect of their sensory consciousness.

I'm perfectly willing to doubt myself, though. Maybe I'm wrong and visual experience is broadly crisp and stable. But if so, I'm not the only person who's wrong about this. So also are most of my interlocutors (whom I *hope* I haven't browbeaten too badly) and probably a good number of philosophers and psychologists.<sup>13</sup> We – I, my friends and cobelievers – have erred due to some theory or preconception, perhaps, some blindness, stupidity, oversight, suggestibility. Okay, let's assume that. I need only, now, turn my argument on its head. We *tried* to get it right. We reflected, sincerely, conscientiously, in good faith, at a leisurely pace, in calm circumstances, without external compulsion, and we got it wrong. Introspection failed us. Since what I'm trying to show is the aptitude of introspection to lead to just such errors, that result would only further my overall thesis. Like other skeptical arguments that turn on our capacity for disagreement, it can triumph in partial defeat.

My pessimistic argument does require this, though: People's judgments about their visual experience differ substantially. My interlocutors' opinions about their ongoing visual experience change significantly as a result of their reflections. This mistake in question, whichever side it's on, though perhaps understandable, is large – no miniscule, evanescent detail, no mere subtlety of language. Furthermore, opinions on both sides arise from normal introspective processes – the same types of process, whatever they are, that underwrite most of our introspective claims about consciousness. And finally, the argument requires that those who disagree don't differ in the basic structure of their visual experience in such a way as to mirror precisely their disagreements. Maybe you can successfully attack one of these premises?



In 2002, David Chalmers and David Hoy ran a summer seminar in Santa Cruz, California, for professional philosophers of mind. They dedicated an entire week of the seminar to the “phenomenology of intentionality”, including most centrally the question of whether *thought* has a distinctive experiential character.

There can be little doubt that sometimes when we think, reflect, ruminate, dwell, or ponder, we simultaneously, or nearly so, experience *imagery* of some sort: maybe visual imagery, such as of keys on the kitchen table; maybe auditory imagery, such as hearing silently, in one’s head, “that’s where they are”. Now here’s the question to consider: Does the phenomenology of thinking *consist entirely* of imagery experiences of this sort, perhaps accompanied by feelings such as discomfort, familiarity, confidence? Or does it go beyond such images and feelings? Is there some distinctive phenomenology of thought, additional to or conjoined with the images, perhaps even capable of transpiring without them?

Scholars disagree. Research and reflection generate dissent, not convergence, on this point. This is true historically,<sup>14</sup> and it was true at the Santa Cruz seminar: Polled at week’s end, seventeen participants endorsed the existence of a distinctive phenomenology of thought, while eight disagreed, either disavowing the phenomenology of thought altogether or saying that imagery exhausts it.<sup>15</sup>

If the issue were highly abstract and theoretical, like most philosophy, or if it hung on recondite empirical facts, we might expect such disagreement. But the introspection of current conscious experience – that’s supposed to be *easy*, right?

Thoughts occupied us throughout the week, presumably available to be discerned at any moment, as central to our lives as the seminar table. If introspection can guide us in such matters – if it can guide us, say, at least as reliably as vision – shouldn't we reach agreement about the existence or absence of a phenomenology of thought as easily and straightforwardly as we reach agreement about the presence of the table?

Unless people diverge so enormously that some have a phenomenology of thought and others do not, then someone is quite profoundly mistaken about his own stream of experience. If there is such a thing as a conscious thought, then presumably we have them all the time. How could you go looking for them and simply not find them? Conversely, if there's no distinctive phenomenology of thought, how could you introspect and come to believe that there is – that is, invent a whole category of conscious experiences that simply don't exist? Such fundamental mistakes almost beggar the imagination; they plead of reinterpretation as disagreements only in language or theory, not real disagreements about the phenomenology itself.

I don't think that's how the participants in these disputes see it, though; and, for me at least, the temptation to recast it this way dissipates when I attempt the introspection myself. Think about what you plan to do next, when you're done reading. Now consider: Was there something it was like to have that thought? Set aside any visual or auditory imagery you may have had. The question is, was there something further in your experience, something besides the imagery, something that might qualify as a distinctive phenomenology of thinking? Try it again if you like. Is the answer *so obvious* that you can't imagine someone going wrong about it? Is it as obvious as that your desk has drawers, your shirt is yellow, your shutters cracked? Must disagreements

about such matters *necessarily* be merely linguistic or about philosophical abstracta? Or, as I think, might people genuinely misjudge even this very basic, absolutely fundamental and pervasive aspect of their conscious experience, even after putting their best introspective resources to work?

viii.

In my view, then, we're prone to gross error, even in favorable circumstances of extended reflection, about our ongoing emotional, visual, and cognitive phenomenology. In other chapters of this book I've argued that we're similarly inept in our ordinary judgments about the experience of visual perspective (Chapter 2), of visual imagery (Chapter 3), and of the echoic environment (Chapter 4); and I've raised concerns about our judgments about the doubleness or singleness of visual experience (Chapter 2), about difference tones, afterimages, and subtle illusions (Chapter 5), and about the extent to which sensory experience is sparse or abundant (Chapter 6). The treatment of dreams in Chapter 1 doesn't concern concurrently introspected experience (except in those rare cases where one introspects one's dream experience while dreaming), but fits into the same general pattern. The final chapter, Chapter 8, is more or less a descent into confusion about normal, waking visual experience with one's eyes closed. Maybe I've erred in my interpretation of some of these cases. Still, taken together they are, I think, evidence enough for a generalization: The introspection of current conscious experience, far from being secure, nearly infallible, is faulty, untrustworthy, and misleading, not just *sometimes a little* mistaken, but *massively* mistaken, about a great variety of issues. If

you stop and introspect now, there's likely very little you should confidently say you know about your own current phenomenology. Perhaps the right kind of learning, practice, or care could largely shield us from error – an interesting possibility that merits exploration – but as yet I see no robust scientific support for such hopes.<sup>16</sup>

What about pain, a favorite example for optimists about introspection? Could we be infallible, or at least largely dependable, in reporting ongoing pain experiences? Well, there's a reason optimists like the example of pain – pain and foveal visual experience of a single bright color. It *is* hard, seemingly, to go too badly wrong in introspecting really vivid, canonical pains and foveal colors. But to use *these cases only* as one's inference base rigs the game. And the case of pain is not always as clear as sometimes supposed. There's confusion between mild pains and itches or tingles. There's the football player who sincerely denies he's hurt. There's the difficulty we sometimes feel in locating pains precisely (for the dentist, say) or in describing their character. I see no reason to dismiss out of hand the possibility of genuine introspective error in such cases. Psychosomatic pain, too: Normally we think of psychosomatic pains as genuine *pains*, but is it possible that some, instead, involve sincere belief in pain that doesn't actually exist?

Inner speech – “auditory imagery” as I called it above – can also seem hard to doubt – that I'm silently saying to myself “time for lunch”. But on closer inspection, I find it too slipping from my grasp. I lean toward thinking that there is cognitive phenomenology of the sort described in the previous section, which can transpire without imagery, but as a result I'm not always sure whether some cogitation that seems to be in inner speech is not, instead, imageless. Also: Does inner speech typically involve not just auditory images but motor activity or motor imagery of the vocal apparatus? Is there an

experiential distinction between (more active) inner speaking and (more passive) inner hearing? I almost despair.

Why, then, do people tend to be so confident in their introspective judgments, especially when queried in a casual and trusting way? Here's my guess: Because no one ever scolds us for getting it wrong about our experience and we never see decisive evidence of error, we become cavalier. This lack of corrective feedback encourages of hypertrophy of confidence. Who doesn't enjoy being the sole expert in the room whose word has unchallengeable weight? In such situations, we tend to take up the mantle of authority, exude a blustery confidence – and genuinely *feel* that confidence (what professor doesn't know this feeling?) until we imagine possibly being proven wrong later by another authority or by unfolding events. About our own stream of experience, however, there appears to be no such humbling danger.

ix.

But wait. Suppose I say, "I'm thinking of a pink elephant" – or even, simply, "I'm thinking". I'm sincere, and there's no linguistic mistake. Aren't claims of this sort necessarily self-verifying? Doesn't merely thinking such thoughts or reaching such judgments, aloud or silently, guarantee their truth? Aren't, actually, their *truth* conditions just a subset of their *existence* conditions? – and if so, mightn't this help us out somehow in making a case for the trustworthiness of introspection?

I'll grant this: Certain things plausibly follow from the very having of a thought: that I'm thinking, that I exist, that something exists, that my thought has the content it in

fact has. Thus, certain thoughts and judgments will be infallibly true whenever they occur – whatever thoughts and judgments assert the actuality or possibility of the conditions or consequences of having them. But the general accuracy of introspective judgments doesn't follow.

Infallibility is, in fact, cheap. *Anything* that's evaluable as true or false, if it asserts the conditions or consequences of its own existence or has the right self-referential structure, can be infallibly true. The spoken assertion "I'm speaking" or "I'm saying 'blu-bob'" is infallibly true whenever it occurs. The sentence "This sentence has five words" is infallibly true whenever uttered. So is the semaphore assertion "I'm holding two flags". So, sure, certain thoughts are infallibly true – true whenever they occur. This shouldn't surprise us; it's merely an instance of the more general phenomenon of self-fulfillment. It has nothing whatsoever to do with introspection; it implies no perfection in the art of ascertaining what's going on in one's mind. If introspection happens to be the process by which thoughts of this sort sometimes arise, that's merely incidental: Infallibly self-fulfilling thoughts are automatically true whether they arise from introspection, from fallacious reasoning, from evil neurosurgery, quantum accident, stroke, indigestion, divine intervention, or sheer frolicsome confabulation.

And how many introspective judgments, really, *are* infallibly self-fulfilling? "I'm thinking" – okay. "I'm thinking of a pink elephant" – well, maybe, if we're liberal about what qualifies as "thinking of" something (see Hintikka 1962; Burge 1988, 1996). But "I'm not angry", "my emotional phenomenology right now is entirely bodily", "I have a detailed image of the Taj Majal, in which every arch and spire is simultaneously well defined", "my visual experience is all clear and stable one hundred degrees into the

periphery”, “I’m having an imageless thought of a pink elephant” – those are a different matter entirely, I’d say.

And, anyway, I’m not so sure we haven’t changed the topic. Does the thought “I’m thinking” or “I’m thinking of a pink elephant” really express a judgment about one’s *conscious experience* (the introspection of which is the topic of this book)? Philosophers might reasonably take different stands here, but it’s not clear to me that I’m committed to believing *anything*, or anything particular, about my conscious experience in accepting such a judgment. I’m certainly not committed to thinking I have a visual image of a pink elephant, or an “imageless thought” of one, or that the words “pink elephant” are drifting through my mind in inner speech. I might hold “I’m thinking of a pink elephant” to be true while I suspect any or all of the latter to be false. Am I committed at least to the view that I’m *conscious*? Maybe. Maybe this is one fact about our conscious experience that we infallibly know. (Could I reach the judgment that I’m conscious *nonconsciously*?<sup>17</sup>) But your ambitions for introspection must be modest indeed if that satisfies you.

x.

I sometimes hear the following objection: When we make claims about our phenomenology, we’re making claims about how things *appear* to us, not about how anything actually *is*. The claims, thus divorced from reality, can’t be false; and if they’re true, they’re true in a peculiar way that permits no error. In looking at an illusion, for example, I may well be wrong if I say the top line is longer; but if I say only that it

*appears* or *seems* longer, I can't in the same way be wrong. The sincerity of the latter claim guarantees its truth. It's tempting, perhaps, to say this: If something *appears* to appear a certain way, necessarily it appears that way. Therefore, we can't misjudge appearances, which is to say, phenomenology.

This reasoning rests on an equivocation between what we might call an *epistemic* and a *phenomenal* sense of "appears" (or, alternatively, "seems"). Sometimes, we use the phrase "it appears to me that \_\_\_\_\_" simply to express a judgment – a hedged judgment, of a sort – with no phenomenological implications whatsoever. If I say, "It appears to me that the Democrats are headed for defeat", ordinarily I'm merely expressing my opinion about the Democrats' prospects. I'm not attributing to myself any particular phenomenology. I'm not claiming to have an image, say, of defeated Democrats, or to hear the word "defeat" ringing in my head. In contrast, if I'm looking at an illusion in a vision science textbook and I say that the top line "appears" longer, I'm *not* expressing any sort of judgment about the line. I know perfectly well it's not longer. I'm making instead, it seems, a claim about my phenomenology, about my visual experience.<sup>18</sup>

Epistemic uses of "appears" *might* under certain circumstances be infallible in the sense I described in the previous section. Maybe, if we assume that they're sincere and normally caused, their truth conditions will be a subset of their existence conditions – I take no stand on that. (See Moran 2001 and Bar-On 2004.) But *phenomenal* uses of "appears" are by no means similarly infallible. This is evident from the case of weak, nonobvious, or merely purported illusions (as discussed in Chapter 5, section v). Confronted with a perfect cross and told there may be a "horizontal-vertical illusion" in the lengths of the lines, one can feel uncertainty, change one's mind, and make what at



least plausibly seem to be errors about whether one line “looks” or “appears” or “seems” in one’s visual phenomenology to be longer than another. You might, for example, fail to notice – or worry that you might be failing to notice – a real illusion in your experience of the relative lengths of the lines; or you might (perhaps under the influence of a theory) erroneously report a minor illusion that actually isn’t part of your visual experience at all. Why not?

Philosophers who speak of “appearances” or “seemings” in discussing consciousness invite conflation of the epistemic and phenomenal senses of these terms. They thus risk breathing an illegitimate air of indefeasibility into our reflections about phenomenology. “It appears that it appears that such-and-such” may have the look of redundancy, but on disambiguation the redundancy vanishes: “It epistemically seems to me that my phenomenology is such-and-such”. No easy argument renders this statement self-verifying.

xi.

Suppose I’m right about one thing – about something that appears, anyway, hard to deny: that people reach vastly different introspective judgments about their emotional experience, their imagery, their visual experience, their thought. If these judgments are all largely *correct*, people must differ immensely in the structure of their conscious experience.

You might be happy to accept that if the price of denying it is skepticism about the accuracy of introspective judgments. Yet I think there’s good reason to pause.

Human variability, though impressive, usually keeps to certain limits. Feet, for example – some are lean and bony, some fat and square, yet all show a common design; skin on the outside, stout bones at the heel, long bones running through the middle into toes, appropriately arranged nerves and tendons. Only in severe injury or birth defect is it otherwise. Human livers may be larger or smaller, but none is made of rubber or attached to the elbow. Human behavior is wonderfully various, yet we wager our lives daily on the predictability of drivers, and no one shows up to department meetings naked. Should phenomenology prove the exception by varying radically from person to person – some of us experiencing one hundred degrees of visual clarity, some only two degrees, some possessed of a distinctive cognitive phenomenology that accompanies every conscious thought, some not, and so forth – with as little commonality as these diverse self-attributions seem to suggest? Of course, if ocular physiology differed in ways corresponding to the differences in report, or if we found vastly different performances on tests of visual acuity or visual memory, or if some of us possessed higher cognition or sympathetic emotional arousal while others did not, that would be a different matter. But as things are, two people walk into a room, their behavioral differences are subtle, their physiologies essentially the same, and yet phenomenologically they're so alien as to be like different species? Hmm!

Here's another possibility: Maybe people are largely the same *except* when they introspect. Maybe we all have basically the same visual phenomenology most of the time, for example, until we reflect directly on that phenomenology – and then some of us experience one hundred degrees of stable clarity while others experience only two

degrees. Maybe we all have a phenomenology of thought, but introspection amplifies it in some people, dissipates it in others; analogously for imagery, emotions, and so forth.

That view has its attractions. But to work it so as to render our introspective judgments basically trustworthy, one must surrender many things. The view concedes to the skeptic that we know little about ordinary, unintrospected experience since it hobbles the inference from introspected experience to experience in the normal, unreflective mode. It threatens to make a hash of change in introspective opinion: If someone thinks a previous introspective view of hers was mistaken – a fairly common experience among people I interview (see, for example, section vi and Chapter 4) – she must, it seems, generally be *wrong* that her previous view was mistaken. She must, generally, be correct, now, that her experience is one way, and also correct, a few minutes ago, that it was quite another way, without having noticed the intervening change. This seems an awkward coupling of current introspective acumen with profound ignorance of change over time. The view renders foolish whatever uncertainty we may sometimes feel when confronted with what might have seemed to be introspectively difficult tasks (as in sections iv and vii, as well as elsewhere in this book). Why feel uncertainty if the judgment one reaches is bound to be right? It also suggests a number of particular – and I’m inclined to think rather doubtful – empirical commitments (unless consciousness is causally irrelevant): major differences in actual visual acuity while introspecting between those reporting broad clarity and those reporting otherwise, major differences in cognition while introspecting between those reporting a phenomenology of thought and those denying it, and so on. The view also requires an entirely different explanation of why theorists purporting to use “immediate retrospection” (such as James 1890/1981, p. 189; Titchener

1912b, p. 491; Hurlburt 1990, ch. 2) also find vastly divergent results – since immediate retrospection, if successful, postpones the act of introspection until after the conscious experience to be reported, when presumably it won't have been polluted by the introspective act.

Is there some compelling reason to take on all this?

xii.

There are two kinds of unreliability. Something might be unreliable because it often goes wrong or yields the wrong result, or it might be unreliable because it fails to do anything or yield any result at all. A secretary is unreliable in one way if he fouls the job, unreliable in another if he neglects it entirely. A program for delivering stock prices is unreliable in one way if it misquotes, unreliable in another if it crashes. Either way, they can't be depended on to do what they ought.<sup>19</sup> Introspection is unreliable in both ways. Reflection on basic features of ongoing experience leads sometimes to error and sometimes to perplexity or indecision. Which predominates in the examples of this chapter, and of this book generally, is not, I think, a deep matter, but rather a matter of context or temperament. Some introspectors will be prone to glib guesswork, while others doubt. Some contexts – for example a pessimistic essay on introspection – will encourage restraint. But whether the result is error or indecision, introspection will have failed – *if* we suppose that introspection ought to yield trustworthy judgments on the matters at hand.

You might reject that last idea. Maybe we shouldn't expect introspection to reveal (for example) the bodily or nonbodily aspects of emotion, or the presence or absence of a distinctive cognitive phenomenology. It wouldn't, then, tell against the reliability of introspection if such cases baffle us. It doesn't tell against the reliability of a stock quote program if it doesn't describe the weather. A passenger car that overheats going 120 miles per hours isn't thereby unreliable. Maybe I've pushed introspection beyond its proper limits, illegitimately forcing it into failure.

What would be the proper limits of introspection, then? Perhaps we can restrain it, rendering it reliable, by restricting ourselves just to the very easiest of judgments – self-fulfilling thought ascriptions, judgments about foveal colors and severe canonical pains, maybe one or few other cases where we really do seem dependably to get it right. (But what cases are those? Sometimes when I think I've found one, further inquiry proves me wrong. See the brief reflections on inner speech in section ix; I had also thought the experience of foveal shape pretty straightforward until considering the issues about perspective discussed in Chapter 2.) Such restraint hardly seems natural: The scientific or philosophical introspector wants to reach judgments about peripherally experienced color, not just foveally experienced color, about the location and character of pain, not just its presence or absence; and the processes in these various cases are not palpably different. Worse, the restraining move deprives introspection of most of its philosophical and psychological value. We wanted to know – didn't we? – general facts about the stream of experience, like about the structure of visual experience and inner speech, like about the range and types of phenomenology we possess. If introspection is to yield knowledge of such matters, it must be permitted to traverse them; and if it

doesn't dependably yield accurate judgments about them, then for the purposes at hand it's unreliable. Or, alternatively, if one's aim is to advance the traditional philosophical foundationalist project of grounding knowledge of the outside world in knowledge of the *sensory experience* that the outside world produces in us, so too will a narrowly caged introspection not suffice: The foundationalist's introspection had better yield detailed knowledge of our sensory experience, and not just a few coarse judgments about foveal color and the like, to underwrite our detailed knowledge of the outside world. If we care about the reliability of introspection, we should care about its reliability in a broad range of cases, not just in a sliver. It helps none to lop its limbs hoping thereby to prevent its sin.

xiii.

Descartes, I think, had it quite backwards when he said the mind – including especially current conscious experience – was better known than the outside world. The teetering stacks of paper around me, I'm quite sure of. My *visual experience* as I look at those stacks, my *emotional experience* as I contemplate the mess, my *cognitive phenomenology* as I drift in thought, staring at them – of these, I'm much less certain. My experiences flee and scatter as I reflect. I feel unpracticed, poorly equipped with the tools, categories, and skills that might help me dissect them. They are gelatinous, disjointed, swift, shy, changeable. They are at once familiar and alien.

The tomato is stable. My visual experience as I look at the tomato shifts with each saccade, each blink, each observation of a blemish, each alteration of attention, with

the adaptation of my eyes to lighting and color. My thoughts, my images, my itches, my pains – all bound away as I think about them, or remain only as interrupted, theatrical versions of themselves. Nor can I hold them still even as artificial specimens – as I reflect on one aspect of the experience, it alters and grows, or it crumbles. The unattended aspects undergo their own changes too. If outward things were so evasive, they'd also mystify and mislead.

I know better what's in the burrito I'm eating than I know my gustatory experience as I eat it. I know it has cheese. In describing my experience, I resort to saying, vaguely, that the burrito tastes "cheesy", without any very clear idea of what this involves. Maybe, in fact, I'm just – or partly – inferring: The thing has cheese, so I must be having a taste experience of "cheesiness". Maybe also, if I know that the object I'm seeing is evenly red, I'll infer a visual experience of uniform "redness" as I look at it. Or if I know that weeding is unpleasant work, I'll infer a negative emotion as I do it. Indeed, it can make great sense as a general strategy to start with judgments about plain, easily knowable facts of the outside world, then infer to what is stranger and more elusive, our conscious experience of that world.<sup>20</sup> I doubt we can fully disentangle such inferences from more "genuinely introspective" processes.

Descartes thought, or is often portrayed as thinking, that we know our own experience first and most directly and then infer from that to the external world.<sup>21</sup> If that's right – if our judgments about the outside world, to be trustworthy, must be grounded in sound judgments about our experiences – then our epistemic situation is dire indeed. However, I see no reason to accept any such introspective foundationalism. Indeed, I suspect the opposite is nearer the truth: Our judgments about the world to a

large extent drive our judgments about our experience. Properly so, since the former are the more secure.



---

<sup>1</sup> See Descartes 1641/1984, especially page 19, as quoted in the epigraph; Hume 1740/1978, especially I.4.2, p. 190 and 192 and I.4.5, page 232 (though Hume may change his mind his *Enquiries* of 1748/1975, see I.1, p. 13 and I.7, p. 60); Sextus 2<sup>nd</sup> c. CE/1994, especially chapters 7 and 10. Pierre Bayle takes a similar position in the entry on Pyrrho in his *Dictionary* (1702/1734-1738, vol. 4, especially remark B on page 654).

<sup>2</sup> See Zhuangzi, 3<sup>rd</sup> c. BCE/1964, Chapter 2; Montaigne 1580/1948, “Apology for Raymond Sebond”. Sanches’s brief treatment of the understanding of the mind (1581/1988, especially p. 243-245 [57-59]) is at most only a partial exception to this tendency. So also is Unger (1975, III.9), who seems to envision only the possibility of linguistic error about current experience and whose skepticism in this instance seems to turn principally on an extremely demanding criterion for knowledge. Huet (1694/2003) is nicely explicit in extending skepticism to internal matters of ongoing thought, though his examples and arguments differ considerably from mine here.

<sup>3</sup> It’s possible that some phenomenologists in the tradition of Brentano, Husserl, Heidegger, and Merleau-Ponty held that introspection was both important and highly untrustworthy, at least in the untrained, as some philosophers have suggested to me in conversation. I find these authors difficult to interpret. Brentano seems the clearest, however, and it does not appear that he holds this view. In *Psychology from an Empirical Standpoint* (1874/1973) he argues that introspection [*innere Beobachtung*] is impossible or scientifically useless because the act of attentive introspective observation necessarily interferes with the mental state under investigation (a charge that goes back to Comte 1830). In its place he recommends “inner perception” [*innere Wahrnehmung*],

---

which does not, evidently, require attention; and this inner perception, he says, is “infallible and does not admit of doubt” (p. 35). So – to judge by these remarks – in Brentano’s view what is difficult or impossible, introspection, is not important, and what is important, inner perception, is infallible.

<sup>4</sup> Skinner, for example, writes:

In summary, a verbal response to a private stimulus may be maintained in strength through appropriate reinforcement based upon public accompaniments or consequences.... [Thus] we may understand why terms referring to private events have never formed a stable and acceptable vocabulary of reasonably uniform usage.... The answer lies in the process by which “terms are assigned to private events”, a process which we have just analyzed in a rough way in terms of the reinforcement of verbal responses.

None of the conditions that we have examined permits the sharpening of reference which is achieved, in the case of public stimuli, by a precise contingency of reinforcement.... It is, therefore, impossible to establish a rigorous scientific vocabulary for public use, nor can the speaker clearly “know himself” in the sense in which knowing is identified with behaving discriminatively (1945, p. 274-275).

The principal flaw of introspective report, according to Skinner (and Bem following him), is vagueness or instability of vocabulary due to the impossibility of precise differential reinforcement.

<sup>5</sup> Some representative passages:

---

I report “I am in pain now.” Now if we take the view that the latter is a piece of indubitable knowledge, to what period of time does the word “now” refer? Not to the time before I started speaking, for there I am depending on memory, which can be challenged. Not the time after I finish speaking, for then I depend on knowledge of the future, which can be challenged too. The time in question must therefore be the time during which the report is being made. But then it must be remembered that anything we say takes time to say. Suppose, then, that I am at the beginning of my report. My indubitable knowledge that I am in pain can surely embrace only the current instant: it cannot be logically indubitable that I will still be in pain by the time the sentence is finished. Suppose, again, that I am just finishing my sentence. Can I do better than remember what my state was when I began my sentence? (Armstrong 1963, p. 420-421).

Thus, according to Armstrong, in principle memory of projection must be involve in self-knowledge, with the consequence that in principle at least error is possible, if only momentarily. Churchland writes:

[T]he taste-sensation of lime sherbet is only very slightly different from the taste-sensation of orange sherbet, and in blindfold tests people do surprisingly poorly at telling which sensation is which. An orange-expectant subject fed lime sherbet may confidently identify her taste-sensation as being of the kind normally produced by orange sherbet, only

---

to retract the identification immediately upon being given a (blind) taste of the genuinely orange article (1988, p. 77).

See also, recently, Lycan 1996; Shoemaker 1996; Kornblith 1998 (reading with a careful eye to distinguish error about current conscious experience from other sorts of error); Dretske 2000; Jack and Shallice 2001; Nichols and Stich 2003; Goldman 2004, 2006; Horgan, Tienson, and Graham 2005; and most of the essays collected in Gertler, ed., 2003. Gertler (2001), Chalmers (2003), and Horgan and Kriegel (2007) have recently attempted to revive restricted versions of (something like) infallibilism. Chalmers's infallibilism is so restricted that I'm not sure how much useful substance remains. Against Gertler, Horgan, and Kriegel, I offer this chapter as argument. Alston 1971 contains a series of quotes illustrating the widespread acceptance of infallibilism and indubitabilism by leading philosophers from the early modern period through the middle of the 20<sup>th</sup> century.

<sup>6</sup> Contrast, for example, Dennett's seeming skepticism about the accuracy of introspective reports in Chapter 11 of his 1991 book with his apparent claims of infallibility or incorrigibility on pages 81 and 96 of that same book and his much more explicit claim of incorrigibility and near infallibility in his 2002 article. I explore the apparent contradictions in Dennett's view in Schwitzgebel 2007b. See Dennett 2007 for his reply.

<sup>7</sup> Philosophers sometimes say that we introspect our attitudes, such as our beliefs and desires, as well as our conscious experiences (see Schwitzgebel 2009). I disagree. The means by which we learn about our attitudes are very different from the means by which we learn about our stream of conscious experience, and fit less well I think with

---

the intuitive notion of introspection (see Schwitzgebel forthcoming-a). My claims in this chapter, and in this book generally, are limited to the knowledge of the stream of conscious experience, not knowledge of attitudes, though I also develop a pessimistic account of knowledge of our morally most important attitudes in Schwitzgebel forthcoming-b. (Do we ever consciously experience our attitudes? It's not clear that we do – see Carruthers 2009, for example, and the discussion in section vii of this chapter – but if we do, then I'd suggest that we have both introspective and non-introspective means of knowing those attitudes and the introspective means would typically be the less trustworthy.)

<sup>8</sup> Haybron 2008 presents an impressive array of evidence suggesting that we don't know how (un-)happy we are. For example, a 2007 Gallup poll of Americans (cited in Haybron 2008) found that *98% of respondents* with household incomes over \$75,000 reported being happy, while only 2% reported being “not too happy” (the closest thing to a negative option); and it's typical for more than 90% of Americans to describe themselves as happy or satisfied with their lives in polls and questionnaires of all sorts. Given the rates of clinical depression – with about 10% of the population suffering from major depression in any given year (Kessler et al. 1999) – not to mention other psychological troubles, I, like Haybron, find such avowals very difficult to believe.

Moment-by-moment emotional assessments seem likely to be at least somewhat more realistic. The best study of this I'm aware of is Brandstätter 2001, which examined the frequency of positive and negative emotions in a variety of situations, based on people's reports about their emotions in “time sampling diaries” when they were beeped unexpectedly at various intervals. He found negative emotion reports – anger, fear,

---

sadness, stress, exhaustion, etc. – to occur in about one-third of all sampled moments.

But even such measures of this likely involve positive bias due to the self-deception and impression management.

<sup>9</sup> It's a matter of some dispute exactly how strongly to read Descartes on this point (e.g., Rozemond 2004), and in the case of emotion, at least, he does elsewhere seem to assert the possibility of error: "For experience shows that those who are the most strongly agitated by their passions are not those who know them best, and that the passions are to be numbered among the perceptions which the close alliance between the soul and the body renders confused and obscure" (1649/1985, p. 339).

Although my target in this essay is not just infallibilism, let me mention a few other authors, besides Descartes, who make remarks that suggest some sort of general immunity to error, doubt, or correction. Locke: "For a Man cannot conceive himself capable of a greater Certainty, that to know that any *Idea* in his Mind is such, as he perceives it to be" (1690/1975, IV.2, p. 531). Hume: "For since all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear" (1740/1978, I.4.2, p. 190). Lewis: "[T]here could be no doubt about the presented content of experience as such at the time when it is given" (1946, p. 183). Shoemaker: "Among the incorrigible statements are statements about 'private' experiences and mental events, e.g., pain statements, statements about mental images, reports of thoughts, and so on. These are incorrigible in the sense that if a person sincerely asserts such a statement it does not make sense to suppose, and nothing could be accepted as showing, that he is mistaken, i.e., that what he

---

says is false” (1963, p. 215-216). Emotion is rarely chosen as the example, but the statements appear to be general.

<sup>10</sup> I take this argument to be in the spirit of Armstrong 1963. It needn’t require that the phenomenology and the judgment be entirely “distinct existences” in the sense Shoemaker 1996 criticizes, though of course it assumes that the one state is *possible* without the other. The only reason I see to reject such a possibility is a prior commitment to infallibilism.

<sup>11</sup> A published example is Melanie in Hurlburt and Schwitzgebel 2007, p. 88-91 – though it’s not the best example, because visual experience wasn’t the main topic of inquiry.

<sup>12</sup> Oddly, though, people seem much less willing – and I include myself – to concede lack of precision in experienced color outside the foveal area, despite the fact that both shape and color are poorly detected outside the fovea. Whether this reflects a real disparity between our nonfoveal experiences of shape and color – more “filling-in” and implicit inference about the latter, perhaps? – or whether it reflects a more stubborn recidivism of introspection, I don’t know.

<sup>13</sup> Among recent authors, Dennett (1991), O’Regan (1992), Mack and Rock (1998), Rensink, O’Regan, and Clark (2000), and Blackmore (2002) all deny that visual experience involves a broad expanse of stable clarity, though we differ somewhat in our positive views. Some of these authors believe we don’t visually experience what we don’t attend to – a question addressed in Chapter 6.

<sup>14</sup> The British empiricists – most famously, Locke (1690/1975), Berkeley (1710/1965), and Hume (1740/1978) – appear to have believed that conscious thought is

---

always imagistic. So did many later introspective psychologists influenced by them, notably Titchener (1909, 1910), who argued against advocates of “imageless thought”, particularly as advocated by the “Würzburg group” (whose work is reviewed in Humphrey 1951 and Kusch 1999). Recent discussions include Siewert 1998; Horgan and Tienson 2002; Wilson 2003; Pitt 2004; Robinson 2005; Prinz 2007; and Hurlburt and Akhter 2008; Spener forthcoming. The issue goes back at least to Aristotle, who in *De Anima* 431a (4<sup>th</sup> c. BCE/1936) appears to espouse the view that thought is always imagistic. See also my brief discussion of this issue in Chapter 3, section iii.

<sup>15</sup> Related poll results (taken at the end of the seminar) are available online at [consc.net/neh/pollresults.html](http://consc.net/neh/pollresults.html) (accessed August 2009). I am inclined to read the disagreement between the “no phenomenology of thought” and the “imagery exhausts it” camps as a disagreement about terms or concepts rather than about phenomenology – a disagreement about whether having an image should count as “thinking”. However, I see no similar terminological explanation of the central dispute. As I recall (though this number is not recorded on the website), only two participants, Maja Spener and I, said they didn’t know.

<sup>16</sup> On classical introspective training, see Chapter 5. On the possibility of careful interview about randomly sampled experiences, see Hurlburt and Schwitzgebel 2007. Schooler and Schreiber (2004) assess the current scientific situation reasonably, if not quite as pessimistically. Very recently, there has been some promising work on meditation (e.g., Lutz et al. 2007), but much remains to be done to establish the scientific credibility of meditation as a technique for learning about conscious experience.



---

<sup>17</sup> But see Chalmers 1996 and Dretske 2003 on the possibility that we could be experienceless “zombies” without knowing it. Both Chalmers and Dretske think we do know that we are conscious, but that it’s not straightforward to see *how* we know that.

<sup>18</sup> See Chisholm 1957 and Jackson 1977. Naturally, ordinary and philosophical usage of “appears” is rather more complex than this simply portrayal suggests if one examines the details; but I don’t think that affects the basic observation of this section.

<sup>19</sup> Epistemologists often define “reliability” so that only the first type of failure counts as a failure of reliability – for example, Goldman (1986), who calls the second sort of failure a lack of “power”. It’s a semantic issue, but I think ordinary language is on my side.

<sup>20</sup> Titchener thinks this strategy is common among untutored introspectors and repeatedly warns against it as “stimulus error” or “R-error” (Titchener 1901-1905; Boring 1921; see also Chapter 6). This strategy bears some relation to the strategy that “transparency theorists” like Harman (1990), Dretske (1995, 2000), and Tye (2003) think we always use in reaching judgments about our experience (see Chapter 5, note 8).

<sup>21</sup> Whether this is the best interpretation of Descartes, I am uncertain. My impression is that Descartes is not entirely clear on the point, and sympathetic interpretations of him shift with the mood of the times. The view is also associated with Locke (1690/1975) and Russell (1912).