Philosophers tend to assume that we have excellent knowledge of our own current conscious experience or ‘phenomenology’. I argue that our knowledge of one aspect of our experience, the experience of visual imagery, is actually rather poor. Precedent for this position is found among the introspective psychologists of the late nineteenth and early twentieth centuries. Two main arguments are advanced toward the conclusion that our knowledge of our own imagery is poor. First, the reader is asked to form a visual image, and it is expected that answering questions about certain basic features of that experience will be difficult. If so, it seems reasonable to suppose that people could be mistaken about those basic features of their own imagery. Second, it is observed that although people give widely variable reports about their own experiences of visual imagery, differences in report do not systematically correlate with differences on tests of skills that psychologists have often supposed to require visual imagery, such as mental rotation, visual creativity, and visual memory.

The guiding question of this special issue is ‘Is the Visual World a Grand Illusion?’ There are at least two distinct scenarios on which the answer to this question might be yes. The visual world might justly be called a ‘grand illusion’ if the properties we would naïvely attribute to the world on the basis of our visual experience are very different from the properties of the world as it really is, independently of our visual experience. Or we might call the visual world a grand illusion if our visual experiences are significantly at odds with our judgments about them, regardless of how things stand in the world beyond. In the latter case, the visual
world is an illusion in the sense that nothing like the visual world as we naïvely take it to be actually exists; the stream of visual consciousness that flows through us has quite a different character.

In this paper, I will argue for a limited version of the second scenario, but pertaining to our visual imagery experience rather than to ordinary visual perception. In particular, I will argue that normal people in favourable circumstances make gross and enduring errors about the nature of their visual imagery experiences, i.e., that at least some people persistently and radically mistake the phenomenal character of their visual imagery. Perhaps support for this view will also indirectly lend plausibility to the corresponding thesis regarding visual experience in perception.

In taking this position, I stand at odds with much of the philosophical tradition that regards people as having some especially accurate ‘privileged access’ to their own current conscious experience or ‘phenomenology’. In the middle of the twentieth century, especially in the analytic tradition, the dominant view appears to have been that we have infallible or ‘incorrigible’ access to, or are entitled to certainty about, our current conscious experience (for example, Lewis, 1946; Ewing, 1951; Ayer, 1963; Shoemaker, 1963; Rorty, 1970; see also Chalmers, forthcoming, and Gertler, 2001, for restricted contemporary articulations of this view). Although this view has fallen under attack in recent decades, most philosophers seem still to hold to some fairly robust notion of privileged access: Even if we are not infallible judges of our current conscious experience, it is generally assumed that we are nonetheless excellent, and circumstances must be unusually challenging in some respect for us to go seriously wrong. The putative examples of mistakes about conscious experience offered by philosophers have for the most part been limited to marginal or science-fiction cases, or cases of psychological stress or abnormality; or the mistakes have been hypothesized to be made only for a moment or near the threshold of discriminability; or (especially in the many citations of Nisbett & Wilson, 1977) the mistakes have been limited to the causes of our experiences rather than the experiences themselves (for example, Armstrong, 1963; Churchland, 1988; Hill, 1991; Audi, 1993; Kornblith, 1998).

Few philosophers have offered examples, as I aim to do in this paper, of normal people in calm circumstances of quiet attention making radical mistakes about central features of their current conscious experience.²

My position, of course, depends on there being facts about our phenomenology or stream of consciousness about which it is possible to be right or wrong, and in particular that there are such facts about our visual imagery experiences. It also assumes that calling an experience ‘conscious’ is not tantamount to saying that one is aware of that experience in the sense that being aware of something implies that one is right about it. Additionally, I should make clear that the kinds of mistakes I

---

² Dennett (1991) is a possible exception. Some of his examples seem to be interpretable as examples of mistakes about our own phenomenology, yet he also wants to grant that ‘heterophenomenological’ reports are incorrigible descriptions of ‘what it’s like’ for the subject (see also Dennett, 2000). The Churchlands and Hilary Kornblith may also endorse the possibility of gross mistakes about conscious experience even in favourable circumstances; but if so, their examples do not show it, since they tend to be limited in the ways described above.
have in mind are not merely linguistic mistakes, admitted by all parties to this debate, that turn on difficulties of expression or the misconstrual of terms; and that I accept the mainstream view that our access to our own experiences is in some respects fundamentally different in kind from our access to other people’s — though I insist that such difference in kind does not imply any special accuracy.

In some respects, the position I endorse traces back to the introspective psychology of the late nineteenth and early twentieth centuries. It was not uncommon for psychologists in this period to begin their textbooks with cautions about the difficulty of accurate introspection (for example, Külpe, 1893/1895; Titchener, 1900; 1901), or even to challenge each other’s work by suggesting that their opponents failed accurately to introspect (for example, James, 1884; Angell, 1897; Woodworth, 1906; Titchener, 1915). Many felt that introspection was a skill that could improve substantially with dedicated cultivation (for example, in the course plotted by Titchener’s 1902 lab manual) and that the reports of untutored introspectors should be treated warily. With all this I heartily agree. However, many introspective psychologists appear to have put too much faith in their own well-trained introspective capacities, which may have been more subject to bias than they thought (Boring, 1953); and some of them did not distinguish sufficiently between the conscious experiences revealed by introspection and the cognitive processes that we might now take to cause or underlie those experiences (especially in the English language literature: see Danziger, 1980).

Imagery was of central importance to introspective psychologists, and they thought that even well-trained introspectors could make serious mistakes about it. In the ‘imageless thought’ controversy of that period, for example, Külpe, Binet, Stout and their followers claimed that thought was possible without imagery, a view that Titchener and others adamantly denied (for contemporary reviews of this debate, see Angell, 1911; Ogden, 1911). Each side felt that the other was making serious introspective mistakes about the presence or absence of imagery in their thought. Perky (1910) even seemed to show that well-trained introspectors could be tricked into making the most fundamental mistake about imagery, the mistake of confusing it with genuine sensory experience (of a dim, hazy, shimmering light source).

An earlier debate, more familiar to most philosophers, also apparently reflecting fundamental disagreement about the experience of imagery, was the debate between Locke and Berkeley about abstract ideas. Locke seems to have felt that he could form an image of a triangle that is ‘neither oblique, nor rectangle, neither equilateral, equicrural, nor scalenon; but all and none of these at once’ (1689/1975, p. 596). Berkeley denied that he had the capacity to generate any such images (1710/1965). While Berkeley leaves open the possibility that

---

[3] But see Monson & Hurlburt (1993) for an argument that the introspective psychologists agreed in their introspections more than they thought they did.

[4] Since Locke says ‘idea’, not ‘image’, it is possible to interpret him as thinking of the idea of the triangle as non-imagistic. However, the standard view of Locke seems to be that ideas, on his view, are always imagistic (although not always visually imagistic). In any case, James (1890/1981) and Huxley (1895) clearly acknowledge the possibility of images with vague or indeterminate features, so they could substitute for Locke as opponents to Berkeley if necessary.
Locke’s imagery is just very different from his own, it is evident that Berkeley felt the more plausible view to be that his own failure to experience abstract images was universal and that Locke simply got it wrong about his own conscious experience.

I believe that the participants in these historical debates were right to mistrust their opponents’ reports of their imagery experiences: Imagery is difficult to introspect accurately, and even in favourable circumstances people can make large mistakes about their own current, conscious imagery. In arguing for this position, I will focus exclusively on visual imagery, as opposed to auditory, motor, or other types of imagery, since visual imagery has been the most broadly discussed. My argument has three main elements. First, I will ask the reader to reflect on her own experience of visual imagery; it is expected that answering even some fairly basic questions about that experience will be difficult. Second, I will describe the wide variability in narrative descriptions of imagery experiences, variability so extreme as to, I think, plausibly generate suspicions about the accuracy of those descriptions. Third, I will argue that psychologists have generally failed to find reliable relationships between differences in subjective reports of imagery and performance on cognitive tasks that plausibly require imagery. I will conclude by discussing some alternative explanations for the failure to find such relationships, including the possibility that unclear standards of reporting are to blame.

I: Hard Questions about the Experience of Visual Imagery

Let me ask you to reflect, then, on your own phenomenology as you form and maintain a visual image. Form a visual image of some familiar object, such as the front of your house. If you are now visualizing this, you presumably are having a conscious experience of imagery.\(^5\) Let me now ask you some questions about that experience.

How much of the scene are you able vividly to visualize at once? Can you keep the image of your chimney vividly in mind at the same time you vividly imagine (or ‘image’) your front door? Or does the image of your chimney fade as your attention shifts to the door? If there is a focal part of your image, how much detail does it have? How stable is it? Supposing that you are not able to image the entire front of your house with equal clarity at once, does your image gradually fade away toward the periphery, or does it do so abruptly? Is there any imagery at all outside the immediate region of focus? If the image fades gradually away toward the periphery, does one lose colours before shapes? Do the peripheral elements of the image have colour at all before you think to assign a colour to them? Do any parts of the image? If some parts of the image have indeterminate colour before a colour is assigned, how is that indeterminacy experienced — as grey? — or is it not experienced at all? If images gradually fade from the centre and it is

\(^5\) Often, I will call this conscious experience simply the ‘experience’ of imagery or of visual imagery. To my ear, the word ‘experience’ implies consciousness, although I know some do not hear it that way (for example, Carruthers, 1992), allowing for the possibility of ‘unconscious experiences’. In the final section of this paper, I will briefly discuss the possibility of unconscious imagery (whether ‘experienced’ or not).
not a matter of the colour fading, what exactly are the half-faded images like? Are the shapes themselves somehow indeterminate, contra Berkeley? How much is visual imagery experience like the experience of seeing a picture, or having phosphenes, or afterimages, or dreams, or daydreams?

Most people of whom I ask such questions at some point stumble or feel uncertainty. They seem like hard questions — questions one stands a reasonable likelihood of getting wrong, even in circumstances of calm attention. Some readers, I am sure, will feel confident in answering all of these questions, and they may judge that all others should feel the same way. But if you feel as I do the difficulty and potential for error in at least some of these questions — if you think people could easily come to answer them incorrectly — then you are granting the possibility of normal, patient, reflective people in favourable circumstances making significant mistakes about their own current conscious experiences.

Many of these questions play on uncertainty about experience at the periphery of attention. Necessarily, the periphery is elusive: As soon as you move your attention toward the periphery to determine what it is like, the periphery moves somewhere else. Nevertheless, the periphery appears to constitute a significant part, if not the bulk, of our conscious experience (or if not, so much the better for my thesis, since many people attest to conscious awareness of things outside the focus of attention); so if you are wrong about the periphery, you are wrong about a major part of your conscious experience.

It has been objected to me in conversation that my asking such a barrage of questions artificially induces bewilderment in my respondents, either because some of my questions build in false presuppositions (for example, in assuming that images have a focal and peripheral region) or because attention to the questions detracts from the difficult task of maintaining a single, constant image. However, even if a respondent’s confusion is a consequence of the format of my questions, it is nonetheless confusion about his own current conscious experience. That the confusion is artificially induced only undermines my thesis to the extent that it implies that he has not been reflecting on his experience in favourable circumstances. I put forward a large number of questions because not everyone will feel uncertainty about the same things, but I imagine that for any respondent we could cull out the questions he finds most difficult and return to them later at leisure. In my own case, at least, the outcome of such a procedure is not very different from the outcome when I consider a barrage of questions.

I also want to emphasize that I have no difficulty with the view that there is a level of detail beyond which it is inappropriate to ask questions. There may be no determinate answer to the question of how many speckles are on the speckled hen you are now imagining, just as there is no determinate answer to the question of exactly how tall Hamlet is. My questions are meant to generate uncertainty not about the number of bricks in the imagined chimney but about the higher-level questions, such as exactly how much detail the image specifies — whether there is or is not a determinate number of bricks.6

---

[6] I suppose it is possible that there is no determinate fact whether there is a determinate number of bricks or not. This position is, of course, distinct from the position that there is no determinate number of
II: Variability in Narrative Responses to Questions about Imagery

Since people differ substantially in their perceptual and cognitive abilities, it is reasonable to suppose that they will also differ in their visual imagery. However, as I will describe in this section, narrative reports of imagery vary widely even among apparently normal people — so much, I suggest, that we may reasonably be led to doubt the veracity of at least some of those reports. The ancient Chinese philosopher Mengzi says, ‘When someone makes a shoe for a foot he has not seen, I am sure he will not produce a basket’ (Lau, 1970) — for most traits, human variation exists within certain limits of normality.

In the 1870s, Francis Galton (1880; 1907) asked subjects to visualize a scene, such as their breakfast table as they sat down to it in the morning, and to describe various features of their resulting imagery, including its vividness, colour and breadth of field. He formally surveyed several hundred men and boys and gathered anecdotal reports from a variety of sources. This classic collection of narrative reports about imagery has to my knowledge remained unduplicated through the present (no doubt in part due to our current preference for quantifiable and easily replicable measures), and I shall rely on it as my primary evidence of the variability in narrative reports about imagery.7

Galton’s respondents populate the full range from people who claim to have no imagery whatsoever to people who claim to have imagery as vivid and detailed as ordinary vision (or even more so), with a considerable number of apparently normal respondents at each extreme. Here are some quotes from subjects at the high end of the scale: ‘The image that arises in my mind is perfectly clear. . . . I can see in my mind’s eye just as well as if I was beholding the scene with my real eye’ (1880, p. 310); ‘All clear and bright; all the objects seem to me well defined at the same time’ (1880, p. 305); ‘The mental image appears to correspond in all respects with reality. I think it is as clear as the actual scene’ (ibid.). Several respondents claim to be able to visualize an object from more than one brick, and it has not to my knowledge been widely discussed. Such a position might arise from a pragmatic antirealism about visual imagery, on which talk about visual imagery is a useful fiction, and thus should not be abandoned altogether, but on which there is no fact of the matter whether a fiction that posits a determinate number of speckles is more useful than a fiction that does not; or it might arise from a view on which visual images are insufficiently stable to support predications about certain of their features over even the smallest duration of attention; or it may grow from some other motivation. If one accepts some such species of higher-level indeterminism about visual imagery, one might deploy it to explain why my respondents are so often baffled — but this explanation must be handled delicately if it is meant to preserve the view that the introspection of visual imagery is largely accurate, since many people are quite confident in their (diverse, and on this view not determinately true) judgments about their imagery experiences, and since the view may suggest a phenomenology of imagery at odds with what most ordinary people would accept.

[7] One might legitimately doubt the replicability of Galton’s results. For example, it seems unlikely that a majority of scientists would now claim to have no visual imagery (see below and footnote 8). This presents no difficulty for my thesis, and in fact supports it, since if people’s reports about their imagery experiences in one era conflict with those in another, and if it is reasonable to suppose that the underlying experiences themselves are similar, we can conclude that at least some of the reports must be in error. In any case, rich variability in subjective reports of imagery is readily rediscoverable by anyone who takes the time to seek it among his acquaintances, as I have done. You might be surprised by some of the things people will say if you give them free rein for a few minutes with open-ended questions.
angle at once. For example, one of them says ‘My mental field of vision is larger than the normal one. In the former I appear to see everything from some commanding point of view, which at once embraces every object and all sides of every object’ (1880, p. 314). Galton also claims that he knows many cases of persons mentally reading off scores when playing the pianoforte, or manuscript when they are making speeches. One statesman has assured me that a certain hesitation in utterance which he has at times is due to his being plagued by the image of the manuscript speech with its original erasures and corrections. He cannot lay the ghost, and he puzzles in trying to decipher it (1907, p. 67).

Other subjects say: ‘My powers are zero. To my consciousness there is almost no association of memory with objective visual impressions. I recollect the breakfast table, but do not see it’ (1880, p. 306); ‘No power of visualizing’ (ibid.); ‘My impressions are in all respects so dim, vague and transient, that I doubt whether they can reasonably be called images’ (ibid.). William James, who cites Galton favourably and at length in *Principles of Psychology* (1890/1981), claims that his own powers of visual imagery are very feeble, that he ‘can seldom call to mind even a single letter of the alphabet in purely retinal terms. I must trace the letter by running my mental eye over its contour in order that the image of it shall have any distinctness at all’ (p. 708).

One of Galton’s subjects, a scientist, embarks on a critique of Galton’s questionnaire itself:

> These questions presuppose assent to some sort of a proposition regarding the ‘mind’s eye’ and the ‘images’ which it sees. . . . This points to some initial fallacy. . . . It is only by a figure of speech that I can describe my recollection of a scene as a ‘mental image’ which I can ‘see’ with my ‘mind’s eye’. . . . I do not see it . . . any more than a man sees the thousand lines of Sophocles which under due pressure he is ready to repeat (1880, p. 302, ellipses Galton’s).

In fact, Galton says that ‘the great majority of men of science’ with whom he interacted at the start of his investigations ‘protested that mental imagery was unknown to them, and they looked on me as fanciful and fantastic in supposing that the words “mental imagery” really expressed what I believed everybody supposed them to mean’ (1880, p. 302). Since Galton found little such scepticism among non-scientists, and even a willingness to declare their imagery completely distinct and full of detail in the face of sceptical responses by him, Galton concludes that, contrary to what one might have expected, scientists tend to ‘have feeble powers of visual representation’ relative to the rest of the population (1880, p. 304).

---

[8] Jorge Luis Borges describes a similar phenomenon in a fictional character obsessed with a coin he calls a ‘Zahir’: ‘There was a time when I could visualize the obverse, and then the reverse. Now I see them simultaneously. This is not as though the Zahir were crystal, because it is not a matter of one face being superimposed upon another; rather, it is as though my eyesight were spherical, with the Zahir in the center’ (Borges, 1962, p. 163).

[9] In contrast, Isaac & Marks (1994) find that the vividness of visual imagery claimed by physics students is at least as great as that of students in other majors, as measured by Marks’ (1973) Vividness of Visual Imagery Questionnaire (VVIQ). For more on the VVIQ, see section three of this paper.
Although Galton and James assume that these self-reports accurately reflect a surprising variation in the quantity and quality of visual imagery, I think it is not unreasonable to view the reports with a certain degree of suspicion. At least, before accepting the existence of such extreme variability in the visual imagery of normal people, it seems sensible to ask whether self-reported high and low imagers differ significantly in their success on cognitive tasks that are plausibly aided by the use of visual imagery. Accordingly, James R. Angell (1910), in discussing the imagery literature of his time, stresses the importance of looking for correlations between what he calls ‘objective methods’ of measuring imagery, in which success or failure on a task depends on the nature of a subject’s imagery, and ‘subjective methods’ in which a subject reports features of her imagery.10 If the correlation between objective and subjective methods is poor, it is plausible to suppose that the differences in subjective reports are differences in report only, not reflecting real differences in visual imagery experience. And if differences in imaging ability are as vast as they would seem to be from the reports of Galton’s subjects, one should expect vast corresponding differences in tasks requiring the use of imagery — differences like the difference between a prodigy and a normal person, or between a normal person and a person with severe deficiencies. Antecedently, it seems plausible to doubt that such differences will be prevalent in normal populations.11

III: Recent Attempts to Relate Differences in Subjective Report with Performance on Imagery Tasks

A great body of studies comparing subjective and objective measures of visual imagery has been amassed in the last several decades, with largely discouraging results that I will briefly review here. As I suggested at the end of the previous section, I think such results undermine the credibility of subjective reports of imagery experience.

Through the 1970s, tests calling for numerical or categorical self-ratings of visual imagery experience (such as Betts’ [1909] Questionnaire upon Mental Imagery, Sheehan’s [1967] shortened version of that questionnaire, Gordon’s [1949] Test of Visual Imagery Control, and Marks’ [1973] Vividness of Visual Imagery Questionnaire) failed for the most part to correlate with performance on cognitive tasks that researchers had thought plausibly to involve imagery. Early reviews of these questionnaires thus tended to be primarily negative (for example, Ernest, 1977; J. Richardson, 1980), leading Paivio (1986), otherwise a great

---

[10] Despite some concerns about the values and assumptions implicit in the use of the terms ‘objective’ and ‘subjective’ (see Lloyd, 1994; 1995), I will employ the same usage in this matter.

[11] If one looks at the self-descriptions of established prodigies, one does find that they sometimes claim to have detailed visual imagery of the sort that could explain their special talents (see, for example, cases described in Luria, 1965/1968; Stromeyer & Pspotka, 1970; Sacks, 1995). In such cases, the subjective reports have at least a prima facie plausibility, although it is also possible that in some cases the prodigies are confabulating details of their imagery to explain what they know to be unusual performances.
defender of the importance of visual imagery, to declare that 'self-report measures of imagery tend to be uncorrelated with objective performance tests' (p. 117).

More recently, McKelvie (1995) conducted a very detailed review and meta-analysis of the literature on the most widely tested visual imagery questionnaire, Marks' (1973) Vividness of Visual Imagery Questionnaire (VVIQ), which prompts subjects to form visual images and asks them to rank the vividness of those images on a scale from 1 (‘Perfectly clear and as vivid as normal vision’) to 5 (‘No image at all, you only “know” that you are thinking of the object’). Although McKelvie reaches a tentatively positive assessment of the VVIQ, the picture he paints is nonetheless negative in a number of important respects. For example, perhaps the three most obvious sorts of tests on which psychologists historically expected good visualizers to excel (judging from where the bulk of research has been done) are tests involving the spatial transformation of visualized objects, such as mental rotation tasks; tests of visual creativity; and tests of visual memory. McKelvie’s meta-analysis finds no significant relationship between scores on the VVIQ and tests of skill at spatial transformation or mental rotation (even Marks [1999], generally quick to defend the importance of his questionnaire, concedes the latter); he finds no significant relationship between the VVIQ and tests of visual creativity for people of normal IQ. (see also Antonietti et al., 1997); and he finds only spotty relationships between the VVIQ and tests of visual memory. On the other hand, McKelvie finds strong relationships between VVIQ scores and self-reports of imagery on other questionnaires; tests of hypnotic susceptibility (but see Crawford & Allen, 1996; Kogon et al., 1998); tests involving Gestalt completion of incomplete figures; and tests of motor and physiological control (but see Eton et al., 1998). One can guess how visual imagery might be important for these tasks, but it is worrisome that we don’t see significant differences in performance on the more obvious sorts of tasks as well.

Assessing these results, McKelvie concludes that ‘[o]n balance . . . the evidence favours the construct validity of the VVIQ, with a more definitive conclusion awaiting further research’ (p. 93). Supposing we grant this weakly positive assessment of the VVIQ, it is still true to say that researchers have generally failed to find the dramatic performance differences between self-reported high

See also A. Richardson (1994) for a more positive but less thorough review that doesn’t confine itself to the VVIQ (and doesn’t undertake a meta-analysis).

McKelvie does find three studies that suggest a relationship between VVIQ scores and visual creativity for subjects of high IQ. Interestingly, he finds a parallel result for verbal creativity: no relationship to the VVIQ unless subjects are specially selected for high IQ. What to make of these findings is unclear.

However, McKelvie also notes that the relationship between the VVIQ and other visual imagery questionnaires is in some cases no stronger than the relationship between the VVIQ and self-report measures of non-visual imagery. In fact, McKelvie’s meta-analysis finds the VVIQ to correlate more highly with the Vividness of Movement Imagery Questionnaire (Isaac et al., 1986) and the Vividness of Auditory Imagery Questionnaire (Kunzendorf, 1982) (combined mean correlation coefficient 0.677) than with any other self-report measures involving visual imagery (see also Antonietti et al., 1997; Eton et al., 1998) — a fact that, as McKelvie observes, raises concerns about the discriminative validity of the VVIQ. It is somewhat surprising that McKelvie doesn’t make more of this issue, but I will not pursue the matter further here.
and low visualizers that one might have expected on the basis of the differences in narrative report described by Galton, and consequently that the reports of Galton’s subjects remain to a significant extent unjustified.\(^{15}\)

However, concerns might be raised even about McKelvie’s weakly positive assessment. Although reports of correlations between the VVIQ and performance on various cognitive tasks presumably involving imagery have continued to appear since 1995 (Crawford & Allen, 1996; Wallace \textit{et al.}, 1996; Campos & Fernández, 1997; Tomes & Katz, 1997; Campos & Fernández, 1998; Keogh & Markham, 1998; Winograd \textit{et al.}, 1998; Riske \textit{et al.}, 2000; Walczyk & Taylor, 2000), negative findings have also continued to appear with roughly equal frequency (Antonietti \textit{et al.}, 1997; Campos & Pérez, 1997; Campos \textit{et al.}, 1997; Eton \textit{et al.}, 1998; Wilkinson & Hyman, 1998; Antonietti, 1999; Heaps & Nash, 1999; Kunzendorf \textit{et al.}, 2000; Tomes & Katz, 2000). There are multiple reasons, independent of the validity of the VVIQ as a measure of visual imagery, to expect some positive findings. Psychological variables tend to correlate, sometimes robustly, for a whole variety of reasons apart from those hypothesized by the experimenter (Meehl, 1990); and it is widely recognized in psychology that positive findings, whatever their cause, are more likely to be pursued and published than negative findings (Chara [1992] addresses this issue in the VVIQ literature in particular). It is also widely recognized that when the procedure of a study is not wholly straightforward, the results are often influenced by the experimenter’s expectations (Intons-Peterson [1983] has shown experimentally how experimenter biases can influence imagery reports in particular). In other words, a disorganized smattering of reported correlations is what we should expect if the VVIQ does not accurately measure visual imagery; and, apart from the correlations with other self-report measures, that is largely what we find. As Slee (1995) emphasizes in her commentary on McKelvie (1995), research on the VVIQ does not seem to fall into the pattern of finding mostly positive relationships with skills there is good theoretical reason to believe are aided by vivid visual imagery and finding mostly weak or negative relationships with skills there is good theoretical reason not to believe are so aided, as would be expected if the VVIQ were an accurate measure of imagery vividness.

Another widely tested area in which one might hope to find a correlation between subjective report and cognitive test is in eidetic imagery, sometimes popularly (but in the view of many theoreticians inaccurately) referred to as ‘photographic memory’. Eidetic imagery has been characterized as imagery of prior but now absent visual stimulation, in some respects like afterimages, but

\(^{15}\) It should be noted, however, that McKelvie finds less variability in imagery reports than one might expect from reading Galton, McKelvie’s meta-analysis yielding a mean VVIQ score of 2.307 and a standard deviation of 0.692. Demand characteristics of the survey may explain some of this clustering toward the low (vivid) end of the scale. As Ahsen (1990) notes, the survey begins by asking the subject to ‘think of some relative or friend’ and then to ‘consider carefully the picture that comes before your mind’s eye’, the latter phrase implying that a picture-like image will be experienced. However, Galton’s survey employs similar language. It is possible the narrative format of Galton’s questionnaire was more encouraging of extreme responses than are the Likert scales of the VVIQ or that cultural differences explain the apparent decline in variability of self-reports of imagery (see footnote 7).
distinguished from afterimages by being positively coloured, while afterimages have colours complementary of the objects perceived, and by being motionless and scannable, while afterimages move about the visual field as the eye saccades (Haber & Haber, 1964; Jaensch, 1930).  

16 The presence of eidetic imagery is measured primarily by subjective report (although many recent researchers, following Haber & Haber, 1964, also check that direction of gaze corresponds with the relative location of the details being reported) and is attributed primarily to children. Often, eidetic images are described as being very detailed (for example, Allport, 1924; Jaensch, 1930; but see Leask et al., 1969). Early researchers on eidetic imagery sometimes claimed to find a variety of differences between eidetikers and non-eidetikers in personality, perception and cognition, but the methodology was often obscure or inconsistent (for critical reviews, see Allport, 1928; Klüver, 1933; Gray & Gummerman, 1975). For example, Gray and Gummerman (1975) state that depending in part on the methodology of the study, frequency estimates of eidetic imagery among children were extremely variable, spanning the full range from zero to 100%. Later, more careful research begun and inspired by Ralph Haber and his colleagues in the 1960s (Haber & Haber, 1964; Leask et al., 1969) resolved some of these methodological inconsistencies, but at the price of most of the positive results — so much so that in 1979 Haber concedes that ‘extensive research has failed to demonstrate consistent correlates between the presence of eidetic imagery and any cognitive, intellectual, neurological or emotional measure’ (p. 583).

Overall, self-reports of imagery experience tend to relate poorly, or at best unsystematically, to objective measures of visual imagery. Although I have not shown this to be the case for narrative reports of the sort catalogued by Galton, I see no reason for special optimism about that case. Accordingly, I am inclined to draw the general conclusion that differences in imagery reports do not reliably reflect differences in imagery experience.  

18 Besides looking at numerically scaled visual imagery questionnaires and reports of eidetic imagery, it would be useful to look at studies of ‘non-imagers’: people who claim to have no visual imagery whatsoever. However, studies of non-imagers are very difficult to find, despite Galton’s claim (cited above) that the ‘great majority of men of science’ claimed that ‘mental imagery was unknown to them’. Perhaps people who claim to be non-imagers are less common these days (across a number of studies, McKelvie [1995] reports a mean score of approximately 3, the midpoint of the VVIQ’s 5-point scale, for ‘poor visualizers’, typically defined as either bottom half or bottom third) — but unless human cognition has changed significantly since the late nineteenth century, it is hard to see how the percentage of genuine non-imagers could have declined. (See also Faw, 1997.)
my thesis that people can be, and often are, grossly mistaken about their own current visual imagery experiences.

**IV: Alternative Explanations for the Failure of Self-reports to Correlate with Performance on Tests Apparently Requiring Imagery**

Since we cannot directly measure another person’s visual imagery experiences, we must rely on indirect evidence if we are to reach the conclusion that people can be mistaken in their reports about those experiences. Briefly, my strategy has been to support the plausibility of the view that people can be mistaken about their own imagery experiences by, first, inducing the reader into uncertainty about her own visual imagery, second, describing the strikingly wide variability in narrative reports of visual imagery experience, and third, arguing that there is little if any systematic relationship between people’s self-ratings of visual imagery and their performance on cognitive tasks that plausibly employ visual imagery. I will conclude by addressing what seem to me the most plausible challenges to my inference from the lack of relationship between subjective and objective measures of visual imagery to the conclusion that subjects are grossly mistaken about their own imagery experiences.

(1) The methods of eliciting the subjective reports do not capture the aspects of imagery relevant to performance on cognitive tests. One version of this objection is raised by Ahsen (1985; 1986; 1987) who suggests that vividness is irrelevant, or even detrimental, to some of the functions of imagery. This suggestion has substantial plausibility: For example, in rotating an imagined figure to see if it matches another figure on the page, what would seem to matter is the gross morphology of the figure, not its vividness. Still, however well this version of the objection works against vividness measures like the VVIQ and Betts’ (1909) Questionnaire upon Mental Imagery, the problem with subjective measures is broader than that. For example, the most prominent subjective visual imagery measure that does not focus on vividness, Gordon’s (1949) Test of Visual Imagery Control, does not appear to correlate appreciably better than the VVIQ with performance on cognitive tasks such as spatial reasoning and creativity tasks (Ernest, 1977; Hiscock, 1978; J. Richardson, 1980; Lorenz & Neisser, 1985; Antonietti et al., 1997; González et al., 1997). Ahsen’s argument also does not address non-imagers or people with imagery as weak as James claims his to be, since a certain minimal level of detail would seem to be required for success on the cognitive tasks, assuming those tasks genuinely to require imagery. Perhaps there are systemic difficulties with all the major visual imagery questionnaires, but someone who wishes to mount this sort of argument at least owes us an explanation as to why a century of trying hasn’t yielded anything demonstrably better.

---

Hurlburt (personal communication, April 2002) says that several self-described non-imagers reported visual imagery when subjected to his experience sampling technique. On a related note, see Thomas’ (1989) interesting account of the behaviorist John Watson’s shift from claiming that he had vivid visual imagery to claiming that he had none.
(2) Visual imagery is of no use in most of the cognitive tasks that have been studied in relation to subjective reports of visual imagery. The extreme version of this view is that imagery is cognitively epiphenomenal: Some people have powerful, vivid and lifelike imagery and others have no imagery whatsoever, but despite this dramatic difference their cognitive abilities differ very little. In this case, we need attribute no errors in reporting: Galton’s subjects could each be accurately describing his own imagery. That the differences do not show up in cognitive tests is simply a consequence of the irrelevance of imagery to cognition. I do not want to delve into the complex debate on this subject (see Paivio, 1971; 1986; Pylyshyn, 1973; forthcoming; Kosslyn, 1980; 1994; Block, 1981) other than to note the awkwardness of any position that posits a major faculty that would seem to have a fairly obvious range of purposes but in fact has little purpose at all and little effect on behaviour apart from the power to generate reports. The awkwardness is compounded if one takes subjective reports of imagery with the uncritical credence of Galton and James, since subjects will often claim to have used imagery in a particular way to solve a problem. To the extent one wishes to mitigate extreme epiphenomenalism by allowing that visual imagery serves some important general functions, it becomes mysterious why correlations have not been found between measures like the VVIQ and success on any but a disorganized smattering of tasks.

(3) Both self-reported good and poor visualizers use imagery, but only good visualizers experience that imagery consciously. This position is a variation of the second alternative, except that what is epiphenomenal is not the imagery itself but the conscious experience of it. We don’t ordinarily think of imagery as unconscious, but perhaps a suitably functional approach to imagery can give some sense to the idea of an unconscious image (as in Paivio, 1971). However, unless conscious experience in general is epiphenomenal (an unpopular position: recent discussions include Flanagan, 1992; Chalmers, 1996; Nichols & Grantham, 2001), one would expect to see differences between the cognitive performance of people whose imagery is unconscious, or largely so, and that of people whose imagery is more fully conscious. As I have argued, such differences are for the most part not to be found. Locating the top of the scale also creates challenges for the proponent of this alternative. To fully credit subjects’ reports on this view, we would have to take reports of extremely detailed and vivid imagery as the benchmark of fully conscious imagery and assume that every subject has imagery at roughly that level of detail. Otherwise, one must either grant that there are substantial differences in the level of imagery detail between subjects (and then one is stuck back with the original problem of explaining the lack of correlation between subjective measures of imagery and cognitive tests) or grant that the subjects at the top of Galton’s scale have overdescribed their imagery, in which case one has granted just the sort of error for which I am arguing. But if everyone’s imagery has the level of detail described in the most optimistic and extreme reports of imagery, then it is surprising that we don’t all perform substantially better on imagery tasks.
(4) Variation in subjects’ responses to questions about their imagery is due to difficulties in expression rather than misapprehension of their conscious experience. Even granting that differences in imagery reports do not reliably reflect differences in imagery experience, it does not necessarily follow that we are poor introspectors of our own visual imagery, since it is possible that the lack of correlation is a consequence entirely of difficulties of communication — that although subjects know their own imagery experiences quite well, putting that knowledge into words is so difficult that the self-reports are nearly useless.

Consider Marks’ VVIQ again. Marks tops the scale with the phrase ‘perfectly clear and as vivid as normal vision’. What does this mean? At least two distinct problems arise in interpreting this phrase (for similar concerns, see J. Richardson, 1980; Cornoldi, 1995). The first is simply in understanding what it is for a visual image to be vivid. Vividness has been associated with precision of outline, amount of detail, brightness, saliency, ‘liveliness’ and other factors (Cornoldi et al., 1992; McKelvie, 1995). Even if a subject has excellent introspective knowledge of his imagery experience, in assessing the ‘vividness’ of his imagery he may weigh these factors differently than do other subjects. The second difficulty is in comparing clarity and vividness across different types of experience. When I visit the optometrist and she asks me if what I see through one lens is as clear as what I see through another, I feel I understand the question. Since I am comparing one perceptual visual experience to another, what it is to be ‘clear’ remains the same across the cases; but if I am asked to compare the clarity of my vision without glasses to the clarity of an orchestra heard through a wall, the matter is not so straightforward. Although the experience of visual imagery and the experience of visual perception likely have important commonalities, significant differences also exist between the two, which make it unclear what the criteria are for saying that a visual image is as clear and vivid as normal vision.

Also, consider Galton’s sceptical scientist who declares that there is a fallacy in supposing the existence of a ‘mind’s eye’ that ‘sees’ images. If one takes ‘sees’ in a literal sense here, this scientist is surely right: There is no homunculus who literally sees the images you form. Yet there seems also to be a looser or metaphorical sense of ‘seeing’ on which it is permissible to say that we see our visual imagery. One might thus think the difference Galton finds between scientists’ and non-scientists’ reactions to his questions has more to do with the strictness with which they take the word ‘see’ than with any real difference in their imagery or any genuine mistake about the experience of imagery.

It is thus reasonable to suppose that some of the variation between reports is due to the lack of clear standards for vividness or even for the presence or absence of an image. If subjects can accurately compare the vividness of one image to another, that provides additional support for this view, since it suggests that the problem is not in the introspection of the images themselves but in the comparison of those introspected images to some external standard. To develop this point it would be necessary to conduct studies that examine the relationship between a subject’s ratings of particular images’ vividness (or some other
introspectible image property)\textsuperscript{19} and the effectiveness of those images in support of that subject’s performance on particular cognitive tasks — perhaps with special attention to whether a subject’s idiosyncratic ratings predict performance as well as do her typical ratings. Most studies of the relevance of imagery to cognitive performance do not use the subjects’ own ratings of their imagery experience. Exceptions include Bower (1972), whose subjects more accurately recalled word pairs for which they had formed an image they rated as highly vivid than word pairs for which they had formed an image they rated as less vivid, and Walczyk (1995), who finds a significant correlation between how vivid a subject claims her memory image of a particular familiar object to be and the accuracy of that subject’s memory of that object. However, in Walczyk’s case, at least, the direction of causation is unclear: Is the memory accurate because the image is vivid, or is the subject reporting a vivid image because she knows that she has an accurate memory? More extensive research is needed before any firm conclusions can be drawn.

Although the lack of clear standards of subjective report likely accounts for some of the variation between reports, I think it plausible that there is substantial genuine introspective error as well. Not all the concerns I have raised about vividness translate equally well to other descriptions of visual imagery, such as those deployed in Gordon’s (1949) Test of Visual Imagery Control (which simply asks subjects whether they can imagine certain sorts of situations) and measures of eidetic imagery, and these relate no better to objective performance than the VVIQ; nor is all the vast variation among Galton’s respondents confined to variation in vividness. And although there may be some merit to treating Galton’s sceptical scientist as a mere quibbler over words, it seems unlikely that all Galton’s purported non-imagers were quibblers of that sort. Galton clearly is inquiring about their visual imagery: If they were aware of distinct conscious experiences of visual imagery in contemplating their mornings’ breakfast tables, distinguishable from memory that is not visually imagistic, it would have been perverse for them not to mention it, regardless of any disagreement about the phrasing of Galton’s questions. Moreover, some respondents explicitly deny having any such distinctive experience. Perhaps some of Galton’s scientific respondents allowed their reports to be unduly influenced by their theoretical positions, but if we treat their avowals as nonetheless sincere, this possibility in no way undermines my thesis. Mistakes are no less mistakes for having been theoretically informed. In fact, even among non-scientists, implicit or explicit views about the nature of mental imagery may be one of the major sources of introspective error.

Finally, in introspecting my own imagery, I do not find myself in the position of feeling that I know exactly what the experience is like, fumbling only because I am unable to put my knowledge clearly into words or because I am unsure of the

\textsuperscript{19} There is some evidence that bizarre imagery is easier to recall than imagery of ordinary scenes (see Einstein & McDaniel, 1987, for a review), but bizarreness seems to be more a feature of one’s reaction to an image or one’s assessment of the likelihood of the situation depicted in an image than an intrinsic, introspectible feature of the imagery itself.
standards of reporting. Rather, I am unsure about the experience itself, about how much detail is filled in at any particular time, about how narrowly the imagery experience confines itself to what is in the scope of my immediate attention, about how much visual imagery is like visual perception. Given that this feeling of uncertainty is not limited to matters of linguistic formulation, it would be odd to suppose that I and others who can be brought to a similar uncertainty by probative questioning can only make mistakes of linguistic formulation. The introspection of visual imagery feels difficult if one sets to it conscientiously. We should not be surprised if people go easily awry.

V: Conclusion

I have conducted this brief examination of our introspective knowledge of visual imagery to promote the more general thesis that we can be, and often are, grossly mistaken about our own current conscious experiences even in favourable circumstances of quiet attention. I do not take myself to have here established that general thesis, which pertains to all of conscious experience, but only to have provided some considerations favouring it in one domain.

Since at least the time of Descartes, theoreticians have widely assumed that nothing is more secure than level-headed judgments about current conscious experience. A similar assumption informs and pervades our commonsense understanding of the mind, as revealed by the surprise or disbelief commonly exhibited in the face of apparent psychological or philosophical examples of even moderate mistakes about conscious experience, and sometimes by our ordinary conversation. If, as I believe, this widespread assumption turns out to be incorrect, we must abandon not only research paradigms in psychology and consciousness studies that depend too trustingly on introspection, and foundationalist and anti-sceptical views in epistemology that take knowledge of immediate experience as a starting point, but also some of our ordinary assumptions about our knowledge of our own mental lives and what it’s like to be ourselves. Human judgment about anything as fluid, changeable, skittish and chaotic as conscious experience is bound to error and confusion.

References


Bets, G.H. (1909), The Distribution and Functions of Mental Imagery (New York: Teachers College, Columbia).


Bower, G.H. (1972), 'Mental imagery and associative learning', in Cognition in Learning and Memory, ed. L.W. Gregg (New York: John Wiley & Sons).


Dennett, D.C. (1991), Consciousness Explained (Boston: Little, Brown and Co.).


Haber, R.N. (1979), 'Twenty years of haunting eidetic imagery: Where's the ghost?', *Behavioral and Brain Sciences*, 2, pp. 583–629.


Lewis, C.I. (1946), *An Analysis of Knowledge and Valuation* (La Salle, IL: Open Court).


Meehl, P.E. (1990), 'Why summaries of research on psychological theories are often uninterpretable', *Psychological Reports*, 66, pp. 195–244.

Schwitzgebel, E. (forthcoming), ‘Why did we think we dreamed in black and white?’, *Studies in History and Philosophy of Science*.