Jonathan W. Schooler & Charles A. Schreiber

Experience, Meta-consciousness, and the Paradox of Introspection

Abstract: Introspection is paradoxical in that it is simultaneously so compelling yet so elusive. This paradox emerges because although experience itself is indisputable, our ability to explicitly characterize experience is often inadequate. Ultimately, the accuracy of introspective reports depends on individuals' imperfect ability to take stock (i.e., to become meta-conscious) of their experience. Although there is no ideal yardstick for assessing introspection, examination of the degree to which self-reports systematically covary with the environmental, behavioural, and physiological concomitants of experience can help to establish the correspondence between meta-consciousness and experience. We illustrate the viability of such an approach in three domains, imagery, mind-wandering, and hedonic appraisal, identifying both the situations in which introspections appear to be accurate and those in which they seem to diverge from underlying experience. We conclude with a discussion of the various factors (including issues of detection, transformation, and substitution) that may cause meta-consciousness to misrepresent experience.

The paradox of introspection stems from the fact that personal experience corresponds to that which we know best subjectively, yet least empirically. Subjectively speaking, there is nothing we know better than our own introspective experience. As James (1890) observed 'Introspective observation is what we have to rely on first and foremost and always' (p.185). Indeed, as Descartes (1637) noted long ago, nothing seems to offer more ontological certainty than introspection, as nothing is more incontrovertible than one's knowledge of partaking in experience. If you can introspect, then you may be dreaming, you might be a brain in a vat, but there seems absolutely no question but that you are having an experience.

Although subjectively incontrovertible, it is impossible to directly assess the contents of experience, and thus no decisive way to empirically determine when

Correspondence:
Jonathan W. Schooler, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4.

reported introspections accurately vs. inaccurately characterize underlying experience. The empirical elusiveness of subjective experience is powerfully illustrated by Dennett's (1991) claim that qualia itself is an illusion. Dennett's argument builds on the observation that the fact that we believe ourselves to possess qualia does not in itself provide evidence that we necessarily do. Zombies (philosophical constructs who appear and behave exactly like humans but who lack internal experience) would also believe and vehemently assert that they possessed vivid subjective experience. Dennett further argues that because it is impossible to fully pin down what any particular experience is really like, the construct of subjective experience is too vague to be meaningful. (For example, when individuals wear inverted glasses and, after a while, begin to navigate successfully in the world, is their subjective experience reverted or do they simply learn to adjust to an upside down world?) Given the infeasibility of either demonstrating the existence of subjective experience or characterizing what it is actually like, Dennett concludes, 'nobody is conscious … we are all zombies'.[1]

The striking thing about Dennett's critique of subjective experience is that from an empirical perspective he is absolutely right. When relying on the rules of evidence upon which science is based (i.e., empirical independent reportable observations) we simply cannot know that we are conscious at all. Nevertheless, (and this is where we diverge from Dennett) we do. It seems **self-evident** to us (and we expect to many readers) that the existence of subjective experience, though entirely eluding the sources of empirical evidence upon which all other knowledge is based, nevertheless remains an indisputable fact whose certainty is at least on par with, and arguably exceeds, anything else that we know (for similar arguments see Chalmers, 1996; Descartes, 1637; Sellars, 1963).

Given the paradox of introspection (i.e., that experience is subjectively self-evident but empirically inscrutable), it might seem tempting to simply abandon any attempts at empirical investigations of subjective experience. Indeed, this strategy characterizes the bulk of psychological research of the last century. However, it is our view that by confronting the paradox of introspection head on, we can make headway. Specifically, three important considerations emerge from the observation that experience is at once subjectively self-evident and empirically inscrutable.

Epistemologically speaking, the paradox of introspection tells us that there is no way that we will ever be able to empirically know subjective experience in the manner in which we empirically know other things. Unlike other elusive concepts (e.g., distant galaxies, tiny particles) whose empirical investigation evades analysis due to limitations of measurement, subjective experience is, as Dennett's analysis illustrates, empirically non-demonstrable. This inscrutability means that empirical investigations of subjective experience must necessarily be grounded in humility, cognizant that any account is open to alternative interpretations.

---

[1] After making these claims, Dennett suggests that it would 'be an act of desperate intellectual dishonesty to quote this assertion out of context' (p. 406). Readers are therefore dutifully advised to review the context in which these assertions are made and form their own opinion about whether we have done them justice.

Methodologically speaking, the paradox of introspection illustrates how investigation of subjective experience also affords a tool unique to other modes of inquiry. Specifically, we can use the fact that we have experience to evaluate the models of experience that we develop. The seeking of a concordance of the conclusions of empirical investigations with the first-person perspectives of those generating and evaluating theories represents a central element of phenomenological psychology (e.g., Varela, 1996, but see Schooler & Dougal, 1999). And although perhaps rarely explicitly acknowledged, in all domains of psychology researchers' willingness to accept theoretical conclusions is affected by whether they find such conclusions introspectively compelling. Of course, this inevitably means that some subjectivity is introduced into the science of introspection. But, after all, isn't that the point?

Finally, from a theoretical perspective, the paradox of introspection highlights a distinction that has been implicit throughout the foregoing discussion: namely, that between the actual contents of an experience (experiential consciousness) and one's explicit beliefs about the contents of consciousness, here interchangeably referred to as meta-consciousness (Schooler, 2000; 2002a) or meta-awareness (Jack & Shallice, 2001; Schooler, 2001). Empirically speaking, when we ask participants to introspect, all that we can garner is what they are meta-conscious of (i.e., what they *believe* they are experiencing). From such meta-conscious reports we must then attempt to infer what they are actually experiencing, and it is the possibility of dissociations between experiential consciousness and meta-consciousness that poses the problem.

Although the distinction between experiential consciousness and meta-consciousness does not eliminate the paradox of introspection, it does help us to see how a subjectively informed empirical science of introspection might proceed. On the one hand, the distinction when considered in light of the paradox of introspection, illustrates how self-reports, which necessarily must be based on meta-conscious appraisals of experience, can be flawed, even if they are based on experience, which is itself self-evident. On the other hand, it suggests that if empirical observations regarding when self-reports coincide with other probable indicators of experience are considered in the context of humility and informed first-person perspectives, then it may be possible to scientifically advance principled conjectures about the nature of experience.

In the following analysis, we further flesh out the nature of such an approach drawing on recent considerations of how a distinction between experience and meta-awareness can help to resolve some of the inherent difficulties associated with introspection. We then illustrate the value of this approach in three domains where introspection plays a particularly crucial role: imagery, mind- wandering, and emotion.

Implications of Dissociations Between Experience and Meta-Awareness for Introspection

Some of the earliest observations regarding the challenges of introspectively re-describing experience came, not surprisingly, from William James (1890)

who observed that 'We find ourselves in continual error and uncertainty so soon as we are called on to name and class and not merely feel' (p. 191). James compared attempts to reflectively pin down fleeting subjective experiences to that of seizing 'a spinning top to catch its motion, or trying to turn up the gas quickly enough to see how the darkness looks' (p. 244). James recognized that the challenge of introspective re-description stems in part from the limitations of language, noting that 'absence of a special vocabulary for subjective facts hinders the study of all but the very coarsest of them' (p. 195).

As a consequence of the difficulties of introspection recognized by James and more vehemently decried by Watson (1921) and others, the use of introspective evidence within mainstream psychology was largely avoided throughout much of the past century. However, an important breakthrough in the acceptance of self-report data followed Ericsson and Simon's (1980; 1993) seminal work documenting the many conditions under which think-aloud protocols (where participants simply state whatever thoughts cross their mind as they engage in a task) offer valid data that closely correspond to behaviour without notably affecting performance. Ericsson and Simon further articulated the conditions under which think-aloud protocols are likely to be valid (e.g., if they are done concurrently without participants attempting to explain the processes underlying their thoughts) and when they are likely to be invalid or reactive (e.g., if they are retrospective or require participants to make inferences about why they are behaving in a particular way).

Ericsson and Simon's consideration of the possible reactivity of think-aloud protocols, and their documentation of the frequent correspondence between the contents of individuals' think-aloud protocols and task performance, represents an important foundation upon which the present analysis is based. However, in their enthusiasm to reestablish the validity of self-report techniques, they notably de-emphasized the possibility that verbally reporting one's thoughts could force individuals to verbally re-represent inherently non-verbal experiences. Arguably, in so doing, they overlooked the possibility (and evidence) that distortions or omissions can occur even when individuals simply think aloud (Schooler *et al.*, 1993; Lane & Schooler, in press; though see Ericsson, 2002 for a critique of this perspective).

Recently, three independent sets of investigators offered compelling arguments for how discrepancies between the contents of experience and individuals' explicit awareness of those contents can lead to difficulties in self-report data (Jack & Shallice, 2001; Jack & Roepstorff, 2002; Lambie & Marcel, 2002; Schooler, 2000; 2001; 2002b). As Schooler (2002b) observed: 'If, as argued, meta-consciousness requires re-representing the contents of consciousness, then it follows that as with any recoding process, some information may become lost or distorted in the translation' (p. 342). In the context of interpreting failures to recognize emotions Lambie and Marcel similarly observed: 'Some of the reported cases of "lack of emotion experience" ... seem to exhibit the features of what our framework would describe as lack of *awareness* of first order emotion phenomenology' (p. 249). Likewise Jack and Shallice pointed out that

'introspective reports may be seen as the product of two factors: firstly the "raw data" ... and secondly the conceptual framework, or "model". ... The extent to which subjects are correct depends on the validity of their model' (p. 177).

These researchers also overlapped regarding the metric by which dissociations between experience and meta-consciousness are identified, with each emphasizing the importance of dissociations between self-reports and other indirect indices. For example, both Lambie and Marcel (2002) and Schooler (2000; 2002b) suggested that discrepancies between repressors' physiological responses to stress and their self-reports (e.g., Asendorph & Scherer, 1983) may be indicative of repressors' lack of awareness of their own subjective state. Similarly, Jack and Shallice (2001) argued that discrepancies between children's self-reports and their behaviour in working on algebra problems (Siegler & Stern, 1998) indicated that children 'lacked awareness of their own discovery and use of the strategy' (p. 178).

The striking overlap between these three largely independent analyses of the sources of introspective error suggests that the distinction between experience and meta-awareness is an idea whose time has come (Jack, 2004). Nevertheless, there are also some important differences between these approaches. Though recognizing the value of identifying dissociations between self-reports and behavioural measures, Jack and Shallice's (2001) primary solution for resolving discrepancies in introspection involves 'replacing or refining the subjects' model for understanding their own mental states' (p. 178). While we concur with Jack and Shallice's approach to the extent that we recognize (and indeed have found) that enhanced expertise in a domain can allow verbal protocols that are more informative and less reactive (e.g., Melcher& Schooler, 1996; in press), attempts at refining participants' introspective models in an effort to enhance their accuracy is also rife with challenges (Schooler, 2002a). In particular, there is always the risk that the introduction of new mental understandings may alter the very experience participants are attempting to elucidate.

Lambie and Marcel's (2002) approach to exploring discrepancies between phenomenal consciousness and meta-consciousness largely focuses on pathological cases in which individuals show an absence of awareness of experienced emotional states. Though consideration of the implications of introspective difficulties for individuals with pathologies is important, this emphasis overlooks the cases in which dissociations between consciousness and meta-consciousness may be observed with normal individuals.

The approach taken by Schooler (Schooler, 2000; 2002a,b; Schooler et al., 2003; Schooler & Schreiber, in press), and extended here, focuses on dissociations between consciousness and meta-consciousness that may be involved not only with pathological cases but also with more everyday examples. In so doing, it introduces a number of methodologies that are particularly suited for documenting dissociations in everyday contexts including using experience sampling to catch people engaging in conscious states (e.g., mind-wandering) before they catch themselves (e.g., Schooler et al., in press), and investigating the impact of encouraging meta-conscious reflection on performance (e.g., Wilson & Schooler,

1991). This approach also articulates two distinct types of dissociations between consciousness and meta-consciousness: temporal dissociations in which consciousness occurs in the absence of meta-consciousness, and translation dissociations in which meta-consciousness misrepresents underlying experience. Finally, this approach, at least as explicated here, grounds dissociations between consciousness and meta-consciousness within the larger context of the paradox of introspection.

One critical concern raised by the paradox of introspection is that although it may seem reasonable to interpret dissociations between behavioural responses and self-report as evidence of experience without corresponding meta-awareness, because we cannot directly measure experience, such a conclusion must be viewed with caution. This inescapable fact means that we must approach this topic with humility, ever open to alternative accounts. At the same time, there are various strategies that we may be able to employ in determining when self-reports are more vs. less likely to correspond to underlying experience. These include:

(1) Multiple convergences. The more all sources of evidence converge on the same conclusion, the more confidence we can reasonably have in that conclusion. Thus if individuals' self-reports not only correspond to their behaviours, but also sensibly fluctuate with changes in the environment and consistently map onto changes in underlying physiological state, then we can become increasingly confident that individuals' meta-conscious appraisals of their experience do in fact correspond to its actual underlying content.

(2) Multiple divergences. Conversely, when individuals' self-reports consistently diverge from environmental fluctuations, behavioural responses, and physiological states, we must become increasingly concerned that meta-conscious appraisals are failing to adequately capture underlying experience.

(3) Reactivity of self-report. If encouraging individuals to introspect about their experience systematically alters, and particularly if it impairs, performance, then it seems reasonable to speculate that the introspections may be failing to do justice to the underlying experience.

(4) Appeals to introspection. When dissociations between self-report and other possible indices of experience are observed, there are always two possible accounts: (a) the self-reports misrepresented actual experiences, or (b) the alternative measures do not actually correspond to experience. Although resolution of this dilemma must be a matter of conjecture, educated introspections regarding what seems more consistent with first person experience may be useful for adjudicating between these alternatives.

In the following discussion, we illustrate the viability of the above approach in three domains: imagery, mind-wandering, and hedonic appraisal.

Visual Imagery

For many years psychologists squabbled about what to make of the common introspection that the visual images in our 'mind's eye' share similarities with

actual visual experiences. Some researchers argued that these introspections reflected analogue representations that corresponded to visual percepts (e.g., Kosslyn, 1980; Shepard & Cooper, 1982). Others, however, insisted that visual imagery was no different from symbolic propositions, and thus its unique visual properties were illusory (e.g., Anderson, 1978; Plyshyn, 1981). Although the debate was hard fought, ultimately those arguing for the uniqueness of visual imagery carried the day as evidenced by a recent survey of psychologists (Reisberg *et al.*, 2002). At the same time, however, these methods have also revealed important domains in which introspections about imagery can be misleading. We review these two sides in turn.

Evidence for the authenticity of imagery introspections

A number of different paradigms have been used to demonstrate the close correspondence between the nature of introspective imagery reports and the physical properties of the imaged stimulus, as well as the behavioural and physiological responses to the image. Some of the first evidence for the imagery/perception link came from studies demonstrating the correspondence between the manipulation of visual images and physical percepts to which those images corresponded. For example, research on mental rotation involving assessment of the match between two objects, one of which must be mentally rotated in order to be compared with the other, reveals a very precise relationship between the amount of time it takes to make a judgment and the amount of physical rotation it would require to align the items (Shepard & Metzler, 1971). Similarly, the time it takes to scan from one location to another on a visualized map is highly correlated with the actual distances between those locations (Kosslyn, 1983). Behavioural evidence for the close correspondence between imagery and vision came from studies demonstrating that visual and verbal strategies produced distinct patterns of reaction times associated with verifying the relationship between pictures (e.g., a physical + above a physical *) and sentences (e.g., THE PLUS IS ABOVE THE STAR) (Macleod *et al.*, 1978). Neurocognitive evidence for the close correspondence between vision and imagery was provided by the numerous studies demonstrating that engaging in visual imagery activates brain regions associated with vision (e.g., Farah, 1995a,b).

Evidence for failures of introspections involving imagery

Although the general introspection that imagery is like vision is now strongly supported, there are also situations where introspections about imagery appear to fail. For example, in a letter-imaging task Kosslyn, Cave, Provost, and von Gierke (1988) found that although participants believe that letters are imaged all at once, their reaction times indicated that mental images are constructed in a piecemeal fashion. Similarly, Chambers and Reisberg (1992) found that although participants believe that their images are equally dense all over, their ability to use imagery to make determinations about characteristics of a previously encoded reversible figure (duck/rabbit) depends on which part of the

image they viewed as corresponding to the face. This finding suggests that contrary to introspection, image clarity varies as a function of the direction of attention (Reisberg, 1996).

Another source of error in imagery introspection comes from verbal description. A substantial body of evidence reveals that when individuals verbally introspect (i.e., attempt to describe in great detail) complex non-verbal images such as the appearance of a previously seen face, disruption can ensue. For example, Schooler and Engstler-Schooler (1990) found that participants who described in detail the appearance of a previously seen face were markedly worse than controls on a subsequent recognition test. Additional studies demonstrated that this interference, termed 'verbal overshadowing', generalized to another non-verbal stimulus (colours) but not to a verbalizable one (a spoken statement). Moreover, the type of introspection was found to be critical. Whereas verbal description was disruptive, visual introspection (i.e., just imagining the face) had no effect. Since its original demonstration, verbal overshadowing has been replicated in numerous laboratories, albeit not always (for a meta-analysis see Meissner, 2002). It has also been found to generalize to other visual stimuli including non-verbal forms (Brandimonte et al., 1997), maps (Fiore & Schooler, 2002), and the appearance of mushrooms (Melcher & Schooler, in press), and also other modalities such as audition, including both music (Houser et al., 2003) and voices (Perfect et al., 2002), as well as taste (Melcher & Schooler, 1996). Such disruption suggests that verbal introspection fails to adequately capture ineffable experiences, breaking them apart in a manner that makes it difficult to put back together (see Schooler et al., 1997; Schooler, 2002c for additional theoretical discussions of this effect).

Summary and caveats

Examination of the relationship between imagery and introspection illustrates both the successes and the limitations of introspection. Although once called into serious question, the converging evidence demonstrating the close correspondence between images and external stimuli, behaviours, and brain activation have consistently supported the reality of this experience. Despite such evidence, the empirical inscrutability of experience leaves open the possibility that the correspondences between imagery and perception occur because both are experienced exclusively symbolically. While such an account cannot be ruled out, it is here where the potential upside of the paradox of introspection may be useful in suggesting an appeal to readers' first-person experience. How plausible is it that your sensory experience of a red apple is purely symbolic? And if not, given the similarities between your experience of an apple and your image of an apple, could an image of an apple be exclusively symbolic? While clearly cognizant of the pitfalls of appealing to the introspective observations of the reader, we nevertheless suggest that the confluence of behavioural, physiological, and self-report data, when combined with the first-person perspective, makes a strong case for the reality of imagery.

While the present approach illustrates the successes of introspection for revealing certain aspects of imagery experiences, it also highlights its difficulties in other respects. Specifically, discrepancies between individuals' self-reported introspections and their behaviours on a variety of measures indicate that introspections are not always accurate. Moreover, reactive effects of verbal introspections suggest that this is particularly the case when individuals attempt to translate their non-verbal images into words. However, once again some humility is required in interpreting such findings. For example, in characterizing the effects of verbally describing memory for previously encountered non-verbal stimuli, the re-description process might alternatively be characterized as simply a description of the stimulus. Jack and Roepstorff (2002b) propose such an account in arguing that the errors induced by verbal description are not a consequence of introspection per se. While this is clearly an issue for further research, several sources of evidence suggest that verbal description of non-verbal memories do indeed induce reflective processes, including (1) recognition time following verbal description is increased (Brown & Lloyd-Jones, 2003), (2) face recognition performance is both less accurate and longer when individuals apply reflective strategies (Dunning & Stern, 1994), and (3) the disruptive effects of prior verbal description are attenuated if individuals are forced to make quick recognition decisions (Schooler & Engstler-Schooler, 1990). Thus, the evidence is highly suggestive in indicating that analytic introspective processes induced by describing memories can sometimes disrupt holistic non-verbal recognition processes. If so, such evidence indicates that the verbal description processes are not doing adequate justice to the memories to which they are directed.

Mind-wandering

Introspections about mind-wandering pose a unique challenge because mind-wandering is defined precisely by its *absence* of a correspondence with ongoing external events. Consequently, the now substantial body of research devoted to examining mind-wandering has relied largely on individuals' introspective reports to assess when it is occurring (for recent reviews see Giambra, 1995; Schooler *et al.*, in press; Smallwood *et al.*, 2003). How much confidence should we have in these reports? In fact, when we examine the correspondence of reports of mind-wandering with external events likely to prompt it, behaviours likely to be induced by it, and physiological responses likely to be associated with it, we find good reasons to believe participants when they report that their minds have wandered off. Nevertheless, recent research also suggests that participants can regularly fail to notice that their minds have wandered. Such failures again illustrate the importance for introspective analyses of distinguishing between having an experience versus knowing that one is having it.

Evidence for the accuracy of introspective reports of mind-wandering

Although mind-wandering episodes are by definition unrelated to the primary task, it is nevertheless possible to validate such reports by establishing a

relationship between the nature of the primary task and the likelihood of mind-wandering reports. Not surprisingly, some tasks are more likely to prompt mind-wandering than others. For example, mind-wandering is more likely when individuals are involved in non-engaging tasks such as reading easy material (Grodsky & Giambra, 1990–1991) or working on well practiced tasks (Teasedale *et al.*, 1995). Mind-wandering also fluctuates as a function of stress (e.g., Antrobus *et al.*, 1967), the gender of the experimenter (Singer, 1988), and circadian rhythm (Giambra *et al.*, 1988–1989). Additional evidence of the validity of introspections of mind-wandering comes from the finding that reported mind-wandering epi-sodes are associated with decrements in performance on a variety of demanding tasks including generating random sequences (Teasdale *et al.*, 1995), memory retrieval (Smallwood *et al.*, 2003) and reading comprehension (Schooler *et al.*, in press). Introspective reports of mind-wandering are also substantiated by evidence of a correspondence between self-reports of mind-wandering and neurocognitive changes associated with EEG (Cunningham *et al.*, 2000), galvanic skin response, and heartbeat (Smallwood *et al.*, 2004).

Evidence that introspections may overlook mind-wandering episodes

The consistent environmental, behavioural, and physiological correlates of mind-wandering support the validity of individuals' reports when they indicate that their minds wandered off. But how confident can we be that, when they do not make such reports, their minds necessarily remained on task? Episodes of mind-wandering can also be remarkably elusive, as evidenced by the frequency with which virtually everyone reports episodes of mind-wandering during read-ing despite the best of intentions and the fact that their eyes continued to move across the page. The striking thing about catching oneself 'zoning-out' during reading is that had one realized one's mind had wandered one would presumably have either stopped reading or stopped daydreaming. The fact that two mutually incompatible activities continue, strongly suggests a lack of meta-awareness that one's mind has wandered.

Schooler *et al.* (in press) recently reported a mind-wandering during reading paradigm that was explicitly designed to illustrate the occurrence of initially unnoticed mind-wandering episodes (which they termed 'zoning-out'). In this paradigm, participants read dull text on a computer screen and indicated every time they caught themselves zoning-out. They were then asked various questions including whether they had been aware that they had been zoning-out prior to reporting it. In a second condition, participants were additionally probed inter-mittently and asked to indicate if they had been zoning-out at that moment. The results revealed that participants: (1) frequently caught themselves zoning out during reading, (2) were still often caught zoning out by the probes, and (3) fre-quently reported that they had been unaware that they had been zoning out, par-ticularly when they were caught by the probes. These findings demonstrate that individuals frequently lack meta-consciousness of the fact that their minds had wandered, even when they are in a study where they are specifically instructed to be vigilantly attending to mind-wandering.

Summary and caveats

The evidence for the reality of introspective reports of mind-wandering are supported by the convergence of evidence that mind-wandering fluctuates in predictable ways with the environment, is predictive of behaviour, and corresponds to underlying physiological responses. Nevertheless, due to the empirically inaccessible nature of inner experience we must again maintain some humility in interpreting such data. For example, it is possible that individuals' reports of daydreaming episodes are actually retrospective reconstructions. Indeed such an account has been offered for dreams where it has been suggested that dream states are not themselves conscious, but may simply be retrospectively perceived as conscious states at the time of awakening (Dennett, 1981). While such accounts cannot be entirely ruled out, we again would appeal to the reader's first-person perspectives. Given your own experiences with catching your mind-wandering, how plausible does it seem that your experience of the daydream could have been an illusion, constructed only after you caught it? When the behavioural, environmental, and physiological evidence for daydreaming experiences is considered together with what we expect to be the strong first-person introspections of the reader, the case for the experience of mind-wandering is quite strong.

Although the evidence strongly suggests that the occurrence of mind-wandering experiences closely corresponds to individuals' self-reports, the evidence also reveals failures in introspective awareness. Indeed, the finding that individuals frequently fail to report mind-wandering episodes until they are probed strongly suggests that mind-wandering episodes may commonly fail to enter meta-awareness. Here too, however, we must be cognizant of other possible accounts of this finding, stemming from our inevitably indirect access to others' direct experience. It is possible, for example, that it is not that individuals fail to notice that their minds have wandered but rather that they forgot that the goal of reading is comprehension. Such an account cannot be ruled out entirely and is arguably more plausible than the possibility that mind-wandering episodes could be retrospective reconstructions. Nevertheless, this may be another place where it may be useful to consult the reader's first-person perspective. Imagine catching your mind wandering while you are reading an important document. How plausible is it that you could have explicitly known that your mind had wandered off, but forgotten that the goal of reading is to understand? We expect that many readers will share our intuition that a more introspectively compelling account is that readers intermittently fail to take stock of where their minds are currently residing.

Hedonic Appraisal

Hedonic appraisals (i.e., assessment of one's own pleasure or pain) represent yet another domain in which researchers are necessarily reliant on individuals' introspections. As Myers observed, 'By definition, the final judge of someone's subjective well-being is whomever lives inside that person's skin' (Myers, 2000, p. 57). Indeed, as in the other domains that we have reviewed, individuals'

introspective reports regarding their hedonic states are also often closely calibrated with external events, related behaviours, and physiological responses. However, there are also times in which hedonic reports become uncoupled with such indices, suggesting that although individuals' feelings of happiness may be unassailable, their meta-conscious characterizations of these feelings can sometimes be problematic.

Evidence for the accuracy of introspective hedonic reports

One critical source of validation of the authenticity of hedonic introspections comes from the often close correspondence between fluctuations in environmental factors and hedonic reports. For example, online ratings of noise-induced annoyance (Schreiber and Kahneman, 2000) and pressure-induced pain (Ariely, 1998) have both been found to be close approximations of the current physical intensity of the stimulus. With respect to positive experiences, Wilson and Schooler (1991) found that untrained subjects' ratings of how much they liked different strawberry jams were closely aligned with the overall jam quality as assessed by experts. Admittedly, there are some situations in which hedonic reports systematically diverge from physical fluctuations. However, in many of these cases the divergence between patterns of physical stimulation and hedonic experience are systematic and make perfect sense. For example, Kahneman *et al.* (1993) observed that temperature-induced discomfort (i.e., cold-pressor pain) escalated over the course of a minute even while the water temperature was constant. However, this finding is not surprising when one recognizes that, whereas the temperature of the cold water remained constant, the temperature of the submerged hand's tissue continued to drop.

In addition to mirroring fluctuations in environmental changes, introspective hedonic reports also regularly correspond to individuals' behaviours. For example, there is a close correspondence between the pain ratings that individuals report during a cold-pressor task and the likelihood that they will remove their arm from the water (e.g., Keogh & Herdenfeldt, 2002). Similarly, individuals' appraisals of the positiveness of objects is often, although not always (see discussion below), highly correlated with their subsequent consumptive behaviour. Not surprisingly, people are more likely to drink beverages (Wilson *et al.*, 1984) and take home posters (Wilson *et al.*, 1993) that they rate more positively than those they rate less positively. Hedonic appraisals are also often closely linked to a variety of physiological responses. For example, fluctuations in reported pain have been found to closely correspond to changes in systolic blood pressure (Hilgard, 1969) and to be well modeled by latent variables involving a combination of evoked potential, pupil dilation, and galvanic skin response (Donaldson *et al.*, 2003). Physiological measures that fluctuate throughout the spectrum of hedonic experiences are somewhat fewer but nevertheless still exist. For example, whereas the reported unpleasantness of photos is associated with increased startle reflex response, the reported positiveness of pictures is associated with reduced startle reflex (Lang *et al.*, 1990).

Evidence for failures of hedonic introspections

Although individuals' hedonic reports are often validated by their convergence with other measures, there are some situations in which these very same measures suggest potential failures of introspection. For example, the correspondence between reported pain and changes in the physical intensity of aversive stimuli is markedly reduced when individuals are exposed to suggestions, including hypnosis or placebos, that lead them to believe that they should not be experiencing discomfort (e.g., Milling & Breen, 2003). Similarly, whereas Wilson and Schooler (1991) observed a close correspondence between individuals' reported liking of jams and jam quality, as determined by experts, this correlation was reduced to non-significance when participants analysed why they felt the way they did about the jams.

In addition to online measures, a number of studies have found that introspective reports of the retrospective global hedonic value of episodes fail to fully reflect what was actually experienced moment-by-moment (even when the episode is brief and the introspection occurs immediately after the experience ends). Logically, the total pleasure (or discomfort) derived from an episode should reflect the sum of the hedonic values of the moments that comprise the experience, that is, the integral of pleasure (or pain) over time. Kahneman *et al.* (1997) formalized this principle as a normative theory of experienced utility, but studies of the relationship between patterns of momentary experience and retrospective evaluations of episodes have consistently found that choices and introspective global judgments do not conform to this principle. Specifically, this research finds that individuals' retrospective evaluations overemphasize the pleasure or discomfort at the episode's most extreme moment and at its ending, the 'peak-end rule'. Other moments, and often the episode's duration, have little effect on global hedonic assessments. Although the peak-end rule may not entirely exhaust the influences on global hedonic evaluations, the assignment of greater weight to the peak and end moments has been found for evaluations of pleasant and unpleasant videos (Fredrickson & Kahneman, 1993), the discomfort of medical procedures (Redelmeier & Kahneman, 1996), laboratory pain induced by pressure to the finger (Ariely, 1998), chronic arthritis pain experienced over the course of several days (Stone *et al.*, 2000), and noise-induced annoyance (Schreiber & Kahneman, 2000; Ariely & Loewenstein, 2000; Schreiber, 2001). When combined with subjects being little influenced by the duration of an experience, this pattern of evaluation by peak and end can lead to introspective judgments and choices that deviate from any rational or objective account of what was experienced (Kahneman *et al.*, 1993).

A second source of evidence for failures of hedonic introspections comes from studies revealing discrepancies between hedonic appraisals and related behaviours. A variety of variables has been found to contribute to such dissociations. One source of such discrepancies is analysing the reason for hedonic experiences. For example, Wilson *et al.* (1993) found that when individuals analysed why they felt the way they did about various art posters, they were more likely to select posters that they ultimately did not put up on their walls than they were to

make such choices when they did not analyse their feelings. Discrepancies between behaviour and introspection about affect can also result from subliminally presented affective stimuli. For example, Strahan *et al.* (2002) found that individuals who were subliminally presented with sad faces were no more likely than control subjects to report reduced mood. Nevertheless, they were more likely to engage in strategies that are associated with mood repair, such as choosing to listen to uplifting rather than 'edgy' music.

A third source of evidence of failures of hedonic introspection comes from investigations of discrepancies between hedonic reports and physiological measures. Hilgard (1969) observed that whereas hypnosis attenuated the pain associated with keeping one's arm in cold water, it did not attenuate changes in blood pressure that are typically highly correlated with pain in this task. There are also certain distinct populations of individuals who consistently show dissociations between physiological responses and affective introspections. For example, when shown stressful videos individuals identified as 'repressors' report less stress than control subjects, whereas their physiological responses (e.g., galvanic skin response) suggest higher levels of stress (Asendorpf & Scherer, 1983). Similarly, Adams *et al.* (1996) found that homophobes who were shown explicit movies of individuals engaging in homosexual acts reported an absence of sexual arousal, while physically evidencing considerable arousal as measured by penile tumescence. As Baumeister *et al.* (1998) observe, this latter finding raises the paradoxical question of how 'does someone manage to feel sexually turned off when his or her body is exhibiting a strong positive arousal?' Again, the distinction between experiential consciousness and meta-consciousness may offer a potential answer. Accordingly, individuals may experience the arousal but, because of their strong motivation, fail to become meta-aware of that experience. Indeed, as will be argued in the final section, the distinction between consciousness and meta-consciousness may account for the various situations in which individuals' introspections disconnect from the environmental, behavioural, and physiological indices.

Summary and caveats

As in the other domains reviewed, there appears to be compelling empirical evidence for both the successes and failures of hedonic introspections. On the positive side, the consistent manner in which hedonic appraisals map onto environmental fluctuations, behavioural responses, and physiological measures helps to reassure us that these reports do have a foundation in experience. At the same time, we must also remain wary of the possibility of alternative accounts. For example, it seems plausible that at least some of these correspondences are not due to fluctuations in individuals' experiences per se, but rather in their beliefs about the experiences. In all likelihood, individuals have theories about when hedonic states should be better or worse that would also be expected to correspond to reported fluctuations in hedonic valence (e.g., individuals presumably believe that changes in the temperature of extremely cold water should correspond to changes in pain). Nevertheless, we again would appeal to the

reader's first-person perspective as one additional source of evidence to consider. Does it seem plausible that, when individuals report excruciating pain resulting from keeping their arm submerged in cold water for extended periods of time, this simply reflects their theory about what they think they are feeling? While it may not be possible to rule out such an account, given the vividness of the experience of pain, we expect the reader may find such accounts implausible, at least in many cases.

While certain errors in hedonic reports seem untenable, we argue that others seem more likely. Cases in which individuals' hedonic reports systematically diverge from all other measures raise serious doubts about the accuracy of the introspections, and suggest dissociations between experience and meta-consciousness. Here too some caution in interpretation is necessary, as we can never be certain if discrepancies between self-reports and other potential indicators of emotion reflect an unconscious emotion (Berridge & Winkielman, 2003) or a failure for an emotional experience to enter meta-awareness. Nevertheless, when we apply our own first-person perspective in considering, for example, repressors' failure to acknowledge any feelings of arousal, we agree with Lambie and Marcel (2002) in concluding that 'it is implausible that, for someone who is exhibiting emotional body states and behaviour, that there is "nothing it is like" to be and behave in such a way' (p. 250).

Why Do Introspections Sometimes Fail?

Up to this point our primary goal has been to persuade the reader that by considering how introspections co-vary with other indices it is possible to make a reasonable approximation of when introspections are trustworthy and when they should be viewed with some skepticism. Periodically, we have also suggested that the distinction between consciousness and meta-consciousness may help to conceptualize these dissociations. However, we have largely delayed discussions of the mechanisms that may lead to failures of introspective reports. In this section we turn to this issue.

The basic idea underlying the distinction between consciousness and meta-consciousness is simply that individuals often have experiences without necessarily explicitly introspecting about them. As a consequence introspection can fail for two very general reasons. First, introspection may not be invoked. Such introspective failures are what Schooler (2002b) refers to as 'temporal dissociations' in which experience occurs in the absence of meta-awareness. People's failure to notice that they are zoning-out during reading is a good example of a temporal dissociation.

Second, introspection can fail because, in their attempt to characterize an experience, individuals may distort it. The majority of the introspective difficulties described in this review can be characterized as 'translation dissociations' between consciousness and meta-consciousness. Some of the sources of these translation dissociations may result from processes associated with *detection*, *transformation*, and *substitution*.

Detection

Although many experiences in life are salient, others are more subtle. If an experience is very subtle then it may be difficult to notice its defining elements. Put another way, the 'signal', which introspection must detect and interpret, is weak or ambiguous. This may explain why individuals can fail to notice how they build their mental images (Kosslyn *et al.*, 1998) or that an image is missing a minor detail (Chambers & Reisberg, 1992). Similarly, the several studies that used subliminal presentation of negative stimuli to induce dissociations between hedonic introspections and behaviour may have been successful because their manipulations were so subtle that participants failed to explicitly notice the affective change (e.g., Strahan *et al.*, 2002). Indeed, when a hedonic experience is sufficiently subtle, attempts to 'look' for the positive elements of the experience may actually interfere with our ability to extract those elements (Schooler *et al.*, 2003).

Transformations

Although some introspective tasks appear to be accomplished with relatively little distortion in the presence of an adequate signal, others require intervening operations for which the system is ill-equipped. Both the analysis of perceptual wholes into component features (as when subjects provide detailed descriptions of a face or a flavour) and the synthesis of a set of moments of experience into an episode (as required for global hedonic evaluations) are examples.

One important source of translation dissociations stems from the inherent non-verbalizability of many experiences. If an experience is holistic and ineffable, then attempts at decomposing and translating that experience into words may distort it in systematic ways. This may be at least part of the reason why describing visual images can impair later performance that relies on those images (e.g., Schooler & Engstler-Schooler, 1990). It may also explain why, when individuals attempt to analyse verbally the basis of affective experiences that are inherently non-verbal (e.g., the reason for liking a painting), they can generate appraisals that fail to correspond either with the physical qualities of the stimulus (Wilson & Schooler, 1991) or with participants' subsequent behaviours towards the stimulus (Wilson *et al.*, 1993). In all of these cases, individuals are required to decompose non-verbal holistic experiences into words. This decomposition process may both lead to the generation of specific inappropriate content (e.g., Schooler & Engstler-Schooler, 1990; Meissner *et al.*, 2001) and/or cause individuals to adopt featural/analytical processing styles that are inadequate for dealing with holistic non-verbal experiences (Fallshore & Schooler, 1995; Schooler *et al.*, 1997; Schooler, 2002c).

Just as analysis of experiential wholes can introduce systematic distortions into introspective reports, so too can synthesis of moments that comprise experiences that extend over time. When individuals attempt to take stock of experiences that occurred in the past, they are, of course, vulnerable to the various cognitive distortions associated with recollection. When a judgment requires

aggregating many moments of experience, however, the potential for distortion increases. For example, although memory limitations may introduce error into retrospective judgments of the intensity of a particular moment of noise exposure, forming this judgment is, nevertheless, a far less complex construction than judging the 'total amount of noise' to which one was exposed over some extended time period. In the case of hedonic judgments, due to an inability to compute the integral of moments of pain/pleasure over time, people focus on the most salient aspects of the experience (primarily, the peak and the end) and give little weight to the remaining elements (e.g., sometimes even neglecting the experience's duration).

Substitutions

Introspection may sometimes go awry because the information accessed is not the record of actual experience. Instead of retrieving the memory of conscious experience, people may instead bring beliefs to meta-awareness without realizing that the contents of meta-consciousness may diverge substantially from what was experienced in consciousness. Expectancies and motivations are among the forces involved in this type of error.

First, when individuals have strong expectations about conscious experience they may access the expectation rather than the actual experience. Such an account may, at least partially, explain the impact of hypnosis and placebos (Hilgard, 1978). Accordingly, if individuals genuinely believe that they should not be feeling pain, then they may be less likely to notice it. A similar type of substitution may underlie the translation errors associated with motivation. Under certain circumstances, individuals may be motivated to ignore their subjective state, thereby, avoiding having to explicitly deal with its implications. An unwillingness to acknowledge their experience to themselves may explain why homophobes fail to notice the arousal that they derive from viewing homosexual acts and why repressors fail to report the stress they experience while viewing disturbing movies. In each of these cases, individuals' motivation to avoid acknowledging unacceptable experiences may cause them to ignore those experiences and replace them with more acceptable accounts (see also Schooler, 2001). The distinction between consciousness and meta-consciousness thus offers an alternative threshold for defining the avoidance attributed to repression. Accordingly, rather than banning thoughts and feelings to the unconscious, motivation processes may prevent individuals from noting the occurrence of thoughts and feelings that actually do come to mind. The contents of meta-consciousness, in such cases, may correspond to preferred beliefs rather than actual experience.

Implications for Methodology

In addition to its theoretical value, the identification of specific conditions under which translation dissociations are particularly apt to occur also offers methodological implications regarding techniques for maximizing introspective accuracy. The translation dissociations associated with nonverbalizable experiences

suggest that researchers either (a) focus on more verbalizable experiences (e.g., logical problem solving, Ericsson & Simon, 1980; 1984; 1993), (b) train participants to describe their experiences more effectively (Jack & Shallice, 2001; Melcher & Schooler, in press), (c) teach people mindfulness practices (e.g. Brown & Ryan, 2003), to enable them to become more consistently attuned to their experiences, (d) force quick responses that minimize verbal reflection (Schooler & Engstler-Schooler, 1990), or (e) offer perceptually grounded non-verbal response options that adequately capture the breadth of the experience (e.g., Bartoshuk, 2000). The translation dissociations associated with introspecting about subtle experiences suggest that researchers should be especially wary of such introspections and/or find ways to increase the salience of the critical dimensions. The translation dissociations associated with expectancies suggest that researchers should attempt to identify individuals' expectancies and assess the degree to which fluctuations in expectancy mediate introspective reports. The problems with introspections based on temporal synthesis point to the need for methods to assist people with more fully incorporating retrievable information into their global hedonic evaluations. Finally, the translation dissociations resulting from motivation should be evaluated by determining whether individuals possess a motivation to ignore or misrepresent their experience, and if so, researchers must examine whether differences in this motivation alter self-reports. In all cases, as illustrated in the preceding sections, confidence in introspections can be gained when they are found to systematically co-vary with experience, behaviours, and/or physiological responses. On the other hand, concern is warranted when these independent indices begin to diverge from the introspections to which they should otherwise correspond.

Of course, the fact that we can conceptualize various failures of introspection in terms of dissociations between consciousness and meta-consciousness does not mean that this approach must necessarily explain all such failures. In some cases failures of introspection may simply reflect participants' unwillingness to be forthcoming. For example, repressors might know they are aroused but simply be unwilling to admit it. Fortunately, there are some techniques for exploring such possibilities. For example, individuals have some (albeit imperfect) control over their facial expressions (Ekman, 1991). If repressors did know they were aroused, they would be expected to try to control their facial expressions of arousal, something they have not been observed to do (Asendorpf & Scherer, 1983). Moreover, there are various scales that can help to identify that subset of individuals who may be moderating their responses for self-presentational reasons (e.g., Plant & Devine, 1998).

Even if individuals are not deliberately misleading researchers, their failure to accurately introspect about psychological events does not necessarily imply that those events were experienced in the absence of meta-awareness. Another possibility is that the psychological event never crossed the boundary of consciousness. For example, Berridge & Winkielman (2003) argue that the occurrence of behavioural effects of emotional manipulations (e.g., changes in consumptive behaviour following subliminally presented sad faces) in the absence of any

self-reported changes in emotion implicate 'nonconscious' (i.e., not experienced) emotions. While the notion that emotions can go entirely 'unfelt' is (from our respective first-person perspectives) somewhat difficult to conceive, it cannot be conceptually ruled out. It is extremely difficult to distinguish between an internal event that is experienced, but not explicitly noticed, as opposed to one that is not experienced at all. At a minimum, the distinction between consciousness and meta-consciousness provides an alternative to the standard view that non-reported psychological events must be non-conscious. Ultimately, however, discriminating between whether something was experienced but not meta-conscious, versus not experienced at all, will require the identification of measures that can illuminate the experiential quality of an event. In this regard, it is worth briefly considering how Sperling's (1960) partial report paradigm enhanced our knowledge of the experience of briefly presented images. There was a time when, despite introspections to the contrary, the experimental data suggested that individuals were only conscious of four to five items in a briefly flashed array, as this is all people were able to report (e.g., Morton, 1941). However, using the partial report paradigm (in which participants were cued to report just a portion of a previously presented visual array) Sperling found that people actually experienced the full array, as evidenced by the fact that they were able to report all of the cued section even though they did not know at the time that the array was visible, which portion of it they were going to have to report.

The procedure used by Schooler *et al.* (in press) to explore discrepancies between self-caught and probe-caught episodes of zoning-out during reading may also hold promise for revealing experiences associated with otherwise non-reported psychological events. As with the partial report paradigm, periodic probing of experience enabled participants to notice briefly experienced events (mind-wandering episodes) before they were forgotten. And like the comparisons of the full vs. partial report procedures, comparisons of self-caught vs. probe-caught cognitive events reveal the degree to which experiences can be overlooked if they are not promptly reported. In principle this procedure could be applied to reveal other temporal dissociations of meta-awareness including failures to notice unwanted thoughts (Wegner, 1994) or emotional experiences (Lane, 2000) where experiences may fail to cross the threshold of meta-awareness.

Although to date, no equivalent procedure has been developed to validate translation dissociations, ultimately it seems likely that conceptually related manipulations might be developed. On the assumption that individuals cannot report episodic memories for states that were not experienced at some level, it may be possible to find retrospective evidence for situations in which meta-awareness initially misrepresents experience. For example, if individuals were alerted to the fact that they might have been subliminally exposed to either negative or positive faces on a prior occasion, and/or if they received some type of mindfulness training (e.g. Brown & Ryan, 2003), they might reveal a retrospective sensitivity to changes in their affective state that they did not notice at the time. Evidence that individuals were capable of retrospectively accessing otherwise

unreported states would provide one potential way of documenting that non-reported psychological events were, nevertheless, experienced at some level.

Given the difficulty of distinguishing between psychological events that fail to reach meta-awareness from those that never make it into consciousness, we might reasonably wonder how important this question really is. However, here again is a situation where the paradox of introspection becomes critical. From the third-person perspective of empirical science, what matters is what is measurable, and to the degree that experience itself defies measurement, its relevance might seem moot. However, from a first-person perspective experience is everything. By analogy, from the perspective of the success of a surgery, it may make little difference whether the patient was really unconscious or simply unable to report/remember the experience. However, from the perspective of the patient, this difference matters considerably! Thus despite the inherent difficulty in distinguishing between psychological events that fail to cross the threshold of consciousness versus meta-consciousness, advances on this issue will be critical for revealing the nature of experience both to science and to ourselves.

Acknowledgements

References

Adams, H. E., Wright, L.W. & Lohr, B.A. (1996), 'Is homophobia associated with homosexual arousal?', *Journal of Abnormal Psychology*, **105**, pp. 440–5.

Anderson, J.R. (1978), 'Arguments concerning representations for mental imagery', *Psychological Review*, **85**, pp. 249–77.

Ariely, D. (1998), 'Combining experiences over time: The effects of duration, intensity changes and on-line measurements on retrospective pain evaluations', *Journal of Behavioral Decision-Making*, **11**, pp. 19–45.

Ariely, D. & Loewenstein, G. (2000), 'When does duration matter in judgment and decision making?', *Journal of Experimental Psychology: General*, **129**, pp. 508–23.

Antrobus, J.S., Coleman, R. & Singer, J.L.(1967), 'Signal detection performance by subjects differing in predisposition to daydreaming', *Journal Consulting Psychology*, **31**, pp. 487–91.

Asendorpf, J.B. & Scherer, K.R. (1983), 'The discrepant repressor: Differentiation between low anxiety, high anxiety, and repression of anxiety by autonomic-facial-verbal patterns of behavior', *Journal of Personality and Social Psychology*, **45**, pp. 1334–46.

Bartoshuk, L.M. (2000), 'Psychophysical advances aid the study of genetic variation in taste', *Appetite*, **34**, p. 105.

Baumeister, R.F., Dale, K. & Sommer, K.L. (1998), 'Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial', *Journal of Personality*, **66**, pp. 1081–124.

Berridge, K.C. & Winkielman, P. (2003), 'What is an unconscious emotion? (The case for unconscious "liking")', *Cognition and Emotion*, **17**, pp. 181–211.

Brandimonte, M.A., Schooler, J.W. & Gabbino, P. (1997), 'Attenuating verbal overshadowing through visual retrieval cues', *Journal of Experimental Psychology: Learning Memory and Cognition*, **23**, pp. 915–31.

Brown, C. & Lloyd-Jones, T.J. (2003), 'Verbal overshadowing of multiple face and car recognition: Effects of within- versus across-category verbal descriptions', *Applied Cognitive Psychology*, **17**, pp. 183–201.

Brown, K.W. & Ryan, R.M. (2003), 'The benefits of being present: Mindfulness and its role in psychological well-being', *Journal of Personality and Social Psychology*, **84**, pp. 822–48.

Chalmers, D.J. (1996), *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford University Press).

Chambers, D. & Reisberg, D. (1992), 'What an image depicts depends on what an image means', *Cognitive Psychology*, **24**, pp. 145–74.

Cunningham, S., Scerbo, M.W. & Freeman, F.G. (2000), 'The electrocortical correlates of daydreaming during vigilance tasks', *Journal of Mental Imagery*, **24**, pp. 61–72.

Dennett, D.C. (1981), 'Are dreams experiences?', in D.C. Dennett, *Brainstorms: Philosophical Essays on Mind and Psychology* (Cambridge, MA: MIT Press).

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

Descartes, R. (1637). *Discourse on Method*. Paris.

Donaldson, G.W., Chapman, C.R., Nakamura, Y., Bradshaw, D.H., Jacobson, R.C. & Chapman, C.N. (2003), 'Pain and the defense response: Structural equation modeling reveals a coordinated psychophysiological response to increasing painful stimulation', *Pain*, **102**, pp. 97–108.

Dunning, D. & Stern, L.B. (1994), 'Distinguishing accurate from inaccurate eyewitness identifications via inquiries about decision processes', *Journal of Personality and Social Psychology*, **65**, pp. 818–35.

Ekman, P. (1991), *Telling Lies: Clues to Deceit In the Marketplace, Politics, and Marriage* (New York: W.W. Norton and Co.).

Ericsson, K.A. (2002), 'Towards a procedure for eliciting verbal expression of non-verbal experience without reactivity: Interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis', *Applied Cognitive Psychology*, **16**, pp. 981–7.

Ericsson, K.A. & Simon, H.A. (1980), 'Verbal reports as data', *Psychological Review*, **87**, pp. 215–51.

Ericsson, K.A. & Simon, H.A. (1993), *Protocol Analysis: Verbal Reports As Data* (Cambridge, MA: The MIT Press).

Fallshore, M. & Schooler, J.W. (1995), 'Verbal vulnerability of perceptual expertise', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, pp. 1608–23.

Farah, M.J. (1995a), 'Current issues in neuropsychology of image generation', *Neuropsychologia*, **33**, pp. 1455–71.

Farah, M.J. (1995b), 'The neural bases of mental imagery', in M.S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (Cambridge, MA: MIT Press).

Fiore, S.M. & Schooler, J.W. (2002), 'How did you get here from there: Verbal overshadowing of spatial mental models', *Applied Cognitive Psychology*, **16**, pp. 897–910.

Fredrickson, B.L. & Kahneman, D. (1993), 'Duration neglect in retrospective evaluations of affective episodes', *Journal of Personality and Social Psychology*, **65**, pp. 45–55.

Giambra, L.M. (1995), 'A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought', *Consciousness and Cognition*, **4**, pp. 1–21.

Giambra, L.M., Rosenberg, E.H., Kasper, S., Yee, W. & Sack, D.A. (1988–89), 'A circadian rhythm in the frequency of spontaneous task-unrelated images and thoughts', *Imagination, Cognition, and Personality*, **8**, pp. 307–12.

Grodsky, A. & Giambra, L.M. (1990–91), 'The consistency across vigilance and reading tasks of individual differences in the occurrence of task-unrelated and task-related images and thoughts', *Imagination, Cognition, and Personality*, **10**, pp. 39–52.

Hilgard, E.R. (1969), 'Pain as a puzzle for psychology and physiology', *American Psychologist*, **24**, pp. 103–13.

Hilgard, E.R. (1978), 'Covert pain in hypnotic analgesia: Its reality as tested by the real-simulator design', *Journal of Abnormal Psychology*, **87**, pp. 655–63.

Houser, T., Fiore, S.M. & Schooler, J.W. (2003), 'Verbal overshadowing of music memory: What happens when you describe that tune?' Unpublished manuscript.

Jack, A.I. (2004), 'What is meta-awareness and why does it matter?', Conference presentation, The Brain and Its Self, Washington University, St. Louis, MO.

Jack, A.I. & Roepstorff, A. (2002a), 'Introspection and cognitive brain mapping: From stimulus-response to script-report', *Trends in Cognitive Sciences*, **6**, pp. 333–9.

Jack, A.I. & Roepstorff, A. (2002b), 'The "measurement problem" for experience: Damaging flaw or intriguing puzzle?', *Trends in Cognitive Sciences*, **6**, pp. 372–4.

Jack, A.I. & Shallice, T. (2001), 'Introspective physicalism as an approach to the science of consciousness', *Cognition*, **79**, pp. 161–96.

James, W. (1890), *The Principles of Psychology* (New York: Holt).

Kahneman, D., Fredrickson, B.L., Schreiber, C.A. & Redelmeier, D.A. (1993), 'When more pain is preferred to less: Adding a better end', *Psychological Science*, **4**, pp. 401–5.

Kahneman, D., Wakker, P.P. & Sarin, R. (1997), 'Back to Bentham? Explorations of experienced utility', *Quarterly Journal of Economics*, **112**, pp. 375–405.

Keogh, E. & Herdenfeldt, M. (2002), 'Gender, coping and the perception of pain', *Pain*, **97**, pp. 195–201.

Kosslyn, S.M. (1980), *Image and Mind* (Cambridge, MA: Harvard University Press).

Kosslyn, S.M. (1983), *Ghosts in the Mind's Machine* (New York: W.W. Norton).

Kosslyn, S.M., Cave, C.B., Provost, D.A. & von Gierke, S.M. (1988), 'Sequential processes in image generation', *Cognitive Psychology*, **20**, pp. 319–43.

Lambie, J.A. & Marcel, A.J. (2002), 'Consciousness and the varieties of emotion experience: A theoretical framework', *Psychological Review*, **109**, pp. 219–59.

Lane, R.D. (2000), 'The functional neuroanatomy of affective style', in R.D. Lane & L. Nadel, *Cognitive Neuroscience of Emotion* (New York: Oxford University Press).

Lane, S.M. & Schooler, J.W. (in press), 'Skimming the surface: The verbal overshadowing of analogy retrieval', *Psychological Science*.

Lang, P.J., Bradley, M.M. & Cuthbert, B.N. (1990), 'Emotion, attention, and the startle reflex', *Psychological Review*, 97, pp. 377–95.

MacLeod, C.M., Hunt, E.B. & Mathews, N.N. (1978), 'Individual differences in the verification of sentence-picture relationships', *Journal of Verbal Learning and Verbal Behavior*, **17**, pp. 493–507.

Meissner, C.A. (2002), 'Applied aspects of the instructional bias effect in verbal overshadowing', *Applied Cognitive Psychology*, **16**, pp. 911–28.

Meissner, C.A., Brigham, J.C. & Kelley, C.M. (2001), 'The influence of retrieval processes in verbal overshadowing', *Memory and Cognition*, 29, pp. 176–86.

Melcher, J., & Schooler, J.W. (1996), 'The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste', *The Journal of Memory and Language*, **35**, pp. 231–45.

Melcher, J. M. & Schooler, J.W. (in press), 'Perceptual and conceptual expertise mediate the verbal overshadowing effect in a training paradigm', *Memory and Cognition*.

Milling, L.S. & Breen, A. (2003), 'Mediation and moderation of hypnotic and cognitive- behavioural pain reduction', *Contemporary Hypnosis*, **20**, pp. 81–97.

Morton, N.W. (1941), 'The reciprocity of visual clearness and the span of apprehension', *American Journal of Psychology*, **54**, pp. 553–8.

Myers, D.G. (2000), *The American Paradox: Spiritual Hunger in an Age of Plenty* (New Haven, NJ: Yale University Press).

Perfect, T.J., Hunt, L.J. & Harris, C.M. (2002), 'Verbal overshadowing in voice recognition', *Applied Cognitive Psychology*, **16**, pp. 973–80.

Plant, E.A. & Devine, P.G. (1998), 'Internal and external motivation to respond without prejudice', *Journal of Personality and Social Psychology*, **75**, pp. 811–32.

Pylyshyn, Z.W. (1981), 'The imagery debate: Analogue media versus tacit knowledge', *Psychological Review*, **87**, pp. 16–45.

Redelmeier, D.A. & Kahneman, D. (1996), 'Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures', *Pain*, **66**, pp. 3–8.

Reisberg, D. (1996), 'The non-ambiguity of mental images', in C. Cornoldi, R. Logie, M. Brandimonte, G. Kaufmann, and D. Reisberg, D. (Eds.), *Stretching the Imagination: Representation and Transformation in Mental Imagery* (New York: Oxford University Press).

Reisberg, D., Pearson, D.G. & Kosslyn, S.M. (2002), 'Intuitions and introspections about imagery: The role of imagery experience in shaping an investigator's theoretical views', *Applied Cognitive Psychology*, **17**, pp. 147–60.

Schooler, J.W. (2000), 'Consciousness, meta-consciousness, and the role of self-report', Plenary Address, Towards a Science of Consciousness, Tucson, AZ.

Schooler, J.W. (2001), 'Discovering memories of abuse in the light of meta-awareness', *Journal of Aggression, Maltreatment, and Trauma*, **4**, pp. 105–36.

Schooler, J.W. (2002a), 'Establishing a legitimate relationship with introspection: Response to Jack and Roepstorff', *Trends in Cognitive Sciences*, **6**, pp. 372–3.

Schooler, J.W. (2002b), 'Re-representing consciousness: Dissociations between consciousness and meta-consciousness', *Trends in Cognitive Sciences*, **6**, pp. 339–44.

Schooler, J.W. (2002c), 'Verbal overshadowing produces a transfer inappropriate processing shift', *Applied Cognitive Psychology*, **16**, pp. 989–97.

Schooler, J.W., Ariely, D. & Loewenstein, G. (2003), 'The explicit pursuit and monitoring of happiness can be self-defeating', in J. Carrillo and I. Brocas (Eds.), *Psychology and Economics* (Oxford: Oxford University Press).

Schooler, J.W. & Dougal, S. (1999), 'The symbiosis of subjective and experimental approaches to intuition', *Journal of Consciousness Studies*, **6**, pp. 280–7.

Schooler, J.W. & Engstler-Schooler, T.Y. (1990), 'Verbal overshadowing of visual memories: Some things are better left unsaid', *Cognitive Psychology*, **17**, pp. 36–71.

Schooler, J.W., Fiore, S. M. & Brandimonte, M. (1997), 'At a loss from words: Verbal overshadowing of perceptual memories', in D.L. Medin (Ed.), *The Psychology of Learning and Motivation* (San Diego, CA: Academic Press).

Schooler, J.W., Ohlsson, S. & Brooks, K. (1993), 'Thoughts beyond words: When language overshadows insight', *Journal of Experimental Psychology: General*, **122**, pp. 166–83.

Schooler, J.W., Reichle, E.D. & Halpern, D.V. (in press), 'Zoning-out during reading: Evidence for dissociations between experience and meta-consciousness', in D.T. Levin (Ed), *Thinking and Seeing: Visual Metacognition in Adults and Children* (Cambridge, MA: MIT Press).

Schooler, J.W. & Schreiber, C.A. (in press), 'To know or not to know: Consciousness, meta-consciousness, and motivation', in Forgas, J.P., Williams, K.R. and von Hippel, W. (Eds), *Conscious and Non-Conscious Motivation* (New York: Cambridge University Press).

Schreiber, C.A. (2001), 'Predicting retrospective global hedonic judgments of brief unpleasant episodes: A framework for differentiating the effects of duration, temporal location, and configural features'. Unpublished dissertation.

Schreiber, C.A. & Kahneman, D. (2000), 'Determinants of the remembered utility of aversive sounds', *Journal of Experimental Psychology: General*, **129**, pp. 27–42.

Sellars, W. (1963), *Science, Perception and Reality* (London: Routlede & Kegan Paul).

Shepard, R.N. & Cooper, L.A. (1982), *Mental Images and Their Transformations* (Cambridge, MA: MIT Press).

Shepard, R.N. & Metzler, J. (1971), 'Mental rotation of three-dimensional objects', *Science*, **171**, pp. 701–3.

Siegler, R.S. & Stern, E. (1998), 'Conscious and unconscious strategy discoveries: A microgenetic analysis', *Journal of Experimental Psychology: General*, **127**, pp. 377–97.

Singer, J.L. (1988), 'Sampling ongoing consciousness and emotional experience: Implications for health', in M.J. Horowitz (Ed.), *Psychodynamics and Cognition* (Chicago, IL: University of Chicago Press).

Smallwood, J.M., Baracaia, S.F., Lowe, M. & Obansawin, M. (2003), 'Task unrelated thought whilst encoding information', *Consciousness and Cognition*, **12**, pp. 452–84.

Smallwood, J.M., Obonsawin, M.C. & Heim, S.D. (2003), 'Task unrelated thought: The role of distributed processing', *Consciousness and Cognition*, **12**, pp. 169–89.

Smallwood, J.M., O'Connor, R.C., Sudbery, M.V. & Obonsawin, M. (2004), 'Task unrelated thought and dysphoria: Behavioural and physiological consequences during encoding'. Unpublished manuscript, Glasgow Caledonian University and The University of Strathclyde.

Sperling, G. (1960), 'The information available in brief visual presentations', *Psychological Monographs*, **74**, entire issue.

Strahan, E.J., Spencer, S.J. & Zanna, M.P. (2002), 'Subliminal priming and persuasion: Striking while the iron is hot', *Journal of Experimental Social Psychology*, **38**, pp. 556–68.

Stone, A.A., Broderick, J.E., Kaell, A.T., DelesPaul, P.A.E.G. & Porter, L.E. (2000), 'Does the peak-end phenomenon observed in laboratory pain studies apply to real-world pain in rheumatoid arthritics?' *Journal of Pain*, **1**, pp. 212–17.

Teasdale, J.D., Dritschel, B.H., Taylor, M.J., Proctor, L., Lloyd, C.A., Nimmo-Smith, I., *et al.* (1995), 'Stimulus-independent thought depends on central executive resources', *Memory & Cognition*, **23**, pp. 551–9.

Varela, F.J. (1996), 'Neurophenomenology', *Journal of Consciousness Studies*, **3** (4), pp. 330–49.

Watson, J.B. (1919), *Psychology from the Standpoint of a Behaviorist* (Philadelphia, PA: Lippincott).

Wegner, D.M. (1994), 'Ironic processes of mental control', *Psychological Review*, **101**, pp. 34–52.

Wilson, T.D., Dunn, D.S., Bybee, J.A., Hyman, D.B. & Rotondo, J.A. (1984). Effects of analyzing reasons on attitude-behavior consistency. Journal of Personality and Social Psychology, 47, 5-16.

Wilson, T.D., Lindsey, S. & Schooler, T.Y. (2000), 'A model of dual attitudes', *Psychological Review*, **107**, pp. 101–26.

Wilson, T.D., Lisle, D.J., Schooler, J.W., Hodges, S.D., Klaaren, K.J. & Lafleur, S.J. (1993), 'Introspecting about reasons can reduce post-choice satisfaction', *Personality and Social Psychology Bulletin*, **19**, 331–9.

Wilson, T.D. & Schooler, J.W. (1991), 'Thinking too much: Introspection can reduce the quality of preferences and decisions?', *Journal of Personality and Social Psychology*, **60**, pp. 181–92.