Introspection and cognitive brain mapping: from stimulus–response to script–report

Anthony Ian Jack and Andreas Roepstorff

Cognitive science has wholeheartedly embraced functional brain imaging, but introspective data are still eschewed to the extent that it runs against standard practice to engage in the systematic collection of introspective reports. However, in the case of executive processes associated with prefrontal cortex, imaging has made limited progress, whereas introspective methods have considerable unfulfilled potential. We argue for a re-evaluation of the standard 'cognitive mapping' paradigm, emphasizing the use of retrospective reports alongside behavioural and brain imaging techniques. Using all three sources of evidence can compensate for their individual limitations, and so triangulate on higher cognitive processes.

Anthony Ian Jack*
Depts of Neurology and Radiology, Washington University, Campus Box 8225, 4525 Scott Avenue, St Louis, MO 63110, USA.
*e-mail: ajack@npg.wustl.edu

Andreas Roepstorff
Centre for Functionally Integrated Neuroscience, PET-Centre, Aarhus University, Noerrebrogade 44, DK-8000 Aarhus C, Denmark.

What would be your response if you were told that most cognitive scientists habitually overlook a valuable source of evidence about mental processes? Every time an experiment is conducted, there are data simply waiting to be collected but persistently neglected. You might be even more surprised to discover that no special equipment is needed to access this rich mine of information. For no more than a few hours preparation before an experiment and half an hour spent with each subject, you could gain access to this data. To make a start you need do scarcely more than read the rest of this article. How could this be? The reason is simple: you are already highly familiar with the evidence.

More than a century ago William James [1] wrote: 'Introspective observation is what we have to rely on first and foremost and always. The word introspection need hardly be defined – it means, of course, looking into our own minds.' Introspective observation is not just a pervasive feature of our personal lives. Cognitive scientists use this source of evidence to inform virtually every stage of their work. From the moment we conceive of an experimental paradigm, through piloting and refinement [2], to the interpretation of results, we are guided by considerations of our own experience and the experiences we attribute to others, understood by proxy to our own [3]. The very language of cognitive science is, in substantial part, the language of experience [4]. Discussions are laden with terms that we understand first and foremost by reference to our own internal states: 'consciousness', 'awareness', 'attention', 'recollection', 'perception', 'imagery', 'rehearsal', 'recognition', 'effort', 'dreaming', etc. Many psychological constructs, but by no means all [5], have an agreed 'operational' behavioural definition. Nonetheless, the question of whether the same construct can be applied to other situations is often difficult to determine. Behavioural paradigms can often be formalized in several different ways. Judgements of similarity and difference between paradigms are open to dispute. It is a simple fact that the cognitive characterization of behavioural paradigms ('task analysis') remains a matter of subjective judgement. Further, it is clear that these judgements are frequently, and sometimes explicitly, informed by introspective observation. Discussions of results are frequently sprinkled with hypotheses whose only direct method of verification is introspection [6–8].

Despite the pervasive importance of introspection, the data are rarely collected systematically (but see [9]). Introspective observations are usually only mentioned as throwaway remarks, contributing little more than anecdotal evidence. The only information that most cognitive scientists have about what it is like to do their experiment comes from their own experience of doing the task [2]. The experiences of the actual subjects, whose objective data are by contrast reported at length, are rarely probed at all.

Re-evaluating introspection

Why is introspective evidence treated so cavalierly? One view is that, despite the cognitive revolution, there remains a 'widespread underestimation of the legacy of behaviourism' [10]. There is some truth to this, but the full picture is more complex. Contrary to popular belief, behaviourists such as Watson and Skinner did see verbal reports as a potentially valuable source of evidence [9,11]. More generally, the relegation of introspective methods can be seen as a by-product of a far more pervasive and fundamental intellectual divide between science and the humanities [12]. While psychology has been busy establishing its status as a science, introspection has been appropriated by other intellectual traditions [13]. This split has done little to promote a realistic understanding. Philosophers have held that introspection is 'transparent', meaning that it allows us directly to apprehend the properties of mental states [14]. By contrast, psychologists have treated relatively minor methodological difficulties as though they were insurmountable obstacles [15].

If introspection were a new technology, we might be much quicker to form a balanced view of its strengths and weaknesses. Functional imaging provides an interesting analogy. Initially the images provided by fMRI, for example, appear beguilingly simple – seeming to promise a direct window on the mind. But understanding what these images tell us

 1364-6613/02/$ – see front matter PII: S1364-6613(02)01941-1

Box 1. Subjective reports and objective behaviours

It is frequently observed that much of psychology is based on subject's reports of their perceptual experiences and the contents of their memory. Sometimes this point is raised to defend and promote the more widespread use of introspective reports [a]. At other times, it is used to undermine the notion that any additional efforts to make use of introspective evidence are necessary [b]. To clarify the discussion, we need to recognize two overlapping distinctions: (i) between objective and subjective responses, and (ii) between responses that require the subject to introspect and those that do not. Objective behavioural responses (which are amenable to an information-processing analysis) are either observably true or false, or observably better or worse at achieving a goal. Examples include responses like button presses or reaching movements, which relate to some physical attribute of a stimulus, such as its colour, shape or location. These responses may or may not require the subject to introspect. Subjective responses indicate some aspect of the subject's mental state, and so are not directly verifiable. Again, they do not necessarily require the subject to introspect. Some concurrent verbal protocol procedures rely on the simple verbalization of the contents of immediate memory – an activity that becomes automatic with practice [c]. Retrospective protocols provide the only direct means of verifying whether or not a response has involved introspection.

In general, psychologists have been comfortable with objective responses, and uncomfortable with subjective responses. Frith *et al.* write that reports 'can be right or wrong. Behaviour, on the other hand, simply occurs...We might wrongly interpret its significance, but ideas of truth and falsity do not apply.' [d]. The implication is that reports might simply be false, whereas behaviour has no such problem. However, far less separates the two than this suggests. First, both types of data will be useless if the subject fails to follow the task instructions, or actively attempts to mislead. Second, just as there is a sense in which 'behaviour simply occurs', there is also a sense in which subjects simply cannot be wrong about their own experiential states [e]. The genuine methodological issues, for both types of data, arise when we come to interpret their significance. Subjects certainly can be, and often are, wrong about their mental processes and the causes of their actions [f]. But then, it is no easy matter to figure out what is going on in a person's mind by observing only behaviour, or by collecting brain imaging data.

Everyone can observe behaviour, but only the subject has direct access to experience. Nonetheless, it is the experimenter who must take responsibility for the interpretation of both [g,h]. Subjects should be discouraged from providing reasons or explanations [c,i], and encouraged to provide unembroidered descriptions [g]. When subjects are genuinely describing their experiences, they might pause to recollect them, struggle to find terms to describe them, and qualify their answers. The last two are important because both experimenter and subject must work to establish a common language through which they can understand each other (see Box 2).

References

a Marcel, A.J. (1998) Phenomenal experience and functionalism. In *Consciousness in Contemporary Science*, (Marcel, A.J. and Bisiach, E., eds), pp. 121–158, Clarendon Press

b Boring, E.G. (1953) A history of introspection. *Psychol. Bull.* 50, 169–189

c Ericsson, K.A. and Simon, H.A. (1993) *Protocol Analysis: Verbal Reports as Data* (Revised edn), MIT Press

d Frith, C. *et al.* (1999) The neural correlates of conscious experience: an experimental framework. *Trends Cogn. Sci.* 3, 105–114

e Searle, J.R. (1992) *The Rediscovery of the Mind*, MIT Press

f Nisbett, R.E. and Wilson, T.D. (1977) Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 75, 522–536

g Cytowic, R.E. (1997) Synaesthesia: phenomenology and neuropsychology. In *Synaethesia* (Baron-Cohen, S. and Harrison, J.E., eds), pp. 17–42, Blackwell

h Jack, A.I. (2001) Paradigm lost: a review of *Consciousness Lost and Found* by Lawrence Weiskrantz. *Mind Lang.* 16, 101–107

i Wilson, T.D. and Schooler, J.W. (1991) Thinking too much: introspection can reduce the quality of preferences and decisions. *J. Pers. Soc. Psychol.* 60, 181–192

about cognition requires considerable caution and a great deal of background knowledge. As recently as a decade ago, cognitive scientists held deeply polarized views about the value of functional imaging. Nonetheless, brain imaging is now firmly incorporated into cognitive science's armoury of techniques.

We are not the first to argue for a re-evaluation of introspection. The past thirty years have seen several lucid and valuable treatments of the issue [15–17], yet they have had little influence on scientific practice. Strangely, it might be the very availability of introspective evidence that has slowed the development of a more sophisticated understanding. Instead of facing 'mere' technological challenges, introspection faces difficulties that are conceptual, historic and 'cultural' [18]. Informally, many psychologists express a powerful conviction that introspection is problematic. As justification, they cite well-worn methodological objections [19,20], or provide a highly selective history of difficulties encountered nearly a century ago [21]. They remain ignorant, or do not care, that the debate has moved on considerably [9,11,15–17,22–26].

Science cannot afford to indulge in such lazy prejudice, especially when recent changes in the research agenda provide both opportunity and motive for a reassessment. For the better part of a century, psychology relied predominantly on just one source of evidence – objective behavioural measures (Box 1). Cognitive science has recently forced a wider perspective. As a result, new methods are available for testing and validating introspective observations that were previously dismissed [27]. Furthermore, introspective evidence has never been more relevant to core theoretical issues. Introspective evidence is the only form of evidence that directly bears on consciousness and subjective states [28]. Despite these factors, resistance to it remains. This gives rise to a peculiar anomaly: the subjects of cognitive experiments are almost invariably conscious (Box 2), and yet research on consciousness remains a distinct and largely segregated area of enquiry. Experience is still regarded as a problem, rather than a resource ready to be tapped.

Cognitive mapping

An incentive for a more fundamental 'paradigm-shift' [29] might be emerging. It springs, suitably enough, from the growing realization that functional imaging techniques do not provide a transparent picture of brain function. To understand the problem, it is worth briefly considering the methodology of functional imaging. One simple model, implicit in many published papers, sees imaging investigations as engaged in 'cognitive mapping'. A basic assumption is

Box 2. Second-person scripts and reports

Cognitive scientists study a peculiar type of entity: human subjects. Compared with most other objects of science, humans have two rather unusual properties. They are conscious, and it is possible to communicate with them. These very useful properties allow the experimenter to 'set up' the subject for the experiment simply by telling them what to do. However, consciousness has also presented problems for cognitive science. The standard cognitive psychological experiment sidesteps consciousness by focussing purely on objective measurements, in which 'the subjectivity of the subject' plays no apparent role. This is the so-called third-person perspective, which describes the subject through measurements 'from the outside'. As a response to this, philosophers have argued for the importance of the 'first-person' perspective, a view that focuses on conscious experiences 'seen from the inside'. Bridging the apparently unbridgeable gap between the first-person and the third-person perspectives became a central issue in consciousness studies and philosophy of mind at the end of the 20th century. This was termed 'the hard problem' by Chalmers [a].

The first-person perspective succeeds where cognitive science has failed, in acknowledging conscious experience. However, it creates some problems of its own. Specifically, it raises the issue of how knowledge about conscious experiences can be shared between individuals. The communication of experience involves an exchange. It is not wholly subjective, as objective descriptions are used to establish points of reference. Nor is it wholly objective, because we must draw on our own experience to understand the experiences of others. In parallel with standard linguistic terminology for personal pronouns [b] we call this the 'second-person' perspective [c], because it involves a communicative interaction in which a common frame of reference is established between two or more individuals. This way of interacting with subjects is a well-established method in contemporary fieldwork-based social anthropology (see A. Roepstorff, PhD thesis, Aarhus University, Denmark, 2001). The second-person perspective is adopted both when the experimenter gives the instructions and when the subject describes their experience. When they give instructions, experimenters treat subjects as 'responsible conscious agents' [d] capable of understanding and acting out the 'role' intended. The collection of retrospective reports signifies little more than a continuation of that relationship. Table I places this in a larger context. It shows how the 1st, 2nd and 3rd person perspectives relate to different types of interaction with the subject. Each interaction is associated both with 'intervention in' and with 'representation of' the subject's state [e]. Whereas the first-person perspective involves an *intra*-personal interaction (which may be described as meta-cognition or meta-awareness, see [f]), shared knowledge of experience is always mediated by an *inter*-personal interaction. This second-person nature of reports has methodological implications: the interaction between the experimenter and the subject shapes the response space of the reports – that is, what sort of introspective judgment, description or comparison is appropriate.

References

a Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press
b Benveniste, É. (1966) *Problémes de Linguistique Generale*, Gallimard
c Roepstorff, A. (2001) Brains in scanners: an Umwelt of cognitive neuroscience. *Semiotica* 134, 747–765
d Shallice, T. (1988) Information-processing models of consciousness: possibilities and problems. In *Consciousness in Contemporary Science* (Marcel, A.J. and Bisiach, E., eds), pp. 305–333, Clarendon Press
e Hacking, I. (1982) *Representing and Intervening*, Cambridge University Press
f Schooler, J. (2002) Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends Cogn. Sci.* 6, 339–344

Table I. First, second and third person perspectives on cognitive tasks. Different perspectives correspond to different types of interaction with the subject.

Perspective	Type of interaction	Intervention (independent variable)	Observation (dependent variable)	Primary theoretical domain
1st	With self	Meta-cognition / executive processes	Experience	Phenomenology
2nd	With investigator	Script (instructions)	Report	Social Anthropology
3rd	With environment	Stimulus	Response	Experimental Psychology
3rd	Neurological	TMS/lesion	e.g. fMRI, EEG, electrophysiology	Neuroscience

that the pattern of activity in the brain depends on which cognitive processes are active. To control the cognitive processes being used, the experimenter gives subjects a specific task. Explicit instructions are given to explain the task context and define goals for the subject. In effect, the experimenter provides the subject with a script.

By varying the script and/or the stimuli presented, the experimenter aims to isolate particular cognitive processes, and relate them to observed changes in brain activity. However, this approach depends on two factors. First, variations in the script and/or stimuli must have a consistent effect on cognitive processes, both within and across subjects. Second, the experimenter must correctly identify the cognitive processes elicited by the task. When the cognitive processes under investigation are closely related to the stimuli presented these conditions are relatively easy to meet. However, matters are far less straightforward when higher cognitive processes, which will depend on the exact nature of the task, are involved. First, it becomes critical that subjects have enacted their parts well, and second, it becomes much harder to understand exactly what the role involves, and thus to perform an accurate task analysis.

Task analysis and frontal lobe functions

Nowhere are these problems more evident than in the frontal lobes. Frith writes, 'There was a time when every task seemed to activate dorsolateral prefrontal cortex, and every experimenter was happy to define a different role for this region.' [30]. More dramatically, Duncan and Owen have conducted a systematic review [31], in which they carefully identify paradigms thought to reflect five distinct cognitive processes: response conflict, response to novelty, working-memory capacity, working-memory delay, and perceptual difficulty. They found that activations associated with these processes were practically indistinguishable, involving the same three areas of prefrontal cortex. In their discussion, the authors hint that these areas cannot be functionally differentiated.

However, this view appears premature. Prefrontal regions show considerable diversity in cytoarchitecture and in connectivity with other brain structures [32]. Anatomically distinct regions that lie just a few synapses away from sensory inputs have been functionally differentiated. There is no positive evidence to suggest a lack of functional differentiation in prefrontal cortex. On the other hand, there is every reason to suppose that the functions of areas that lie further from input and output will prove harder to identify. We must at least explore the possibility that functions can be localized within prefrontal cortex by attempting to develop a better set of cognitive constructs.

Several considerations point to the use of introspective information to assist in this development. First, these 'executive' processes are not easily susceptible to standard behavioural analysis. They lie 'some distance upstream of observed behaviours' [33]. Second, it is widely held that frontal activation reflects, and is dependent upon, the goals and strategies adopted by the subject [8,34,35]. As a result, frontal activity is highly dependent on variations in the task instructions and context [36]; that is, on how the subject acts out the 'script' provided by the experimenter. Third, there is already good evidence that introspective evidence can be informative about the cognitive processes associated with prefrontal cortex. Consider the following processes, commonly associated with frontal function:

(1) Working memory [6]: subjects' reports can provide information about rehearsal and chunking strategies [37,38]. They might also allow researchers to develop more accurate and specific labels to replace the quasi-phenomenological terms already in use, such as 'monitoring', 'manipulation', 'selection' and 'maintenance' [36].
(2) Episodic retrieval: a process originally identified by reference to phenomenology [39]. In addition, reports have been used to elaborate a detailed model of processes involved in autobiographical recall [40], which themselves have subsequently been used to inform the interpretation of imaging findings [36].
(3) Episodic encoding: reports have been used to ascertain the number of times information is rehearsed [41], and the use of imagery [42]. In both cases these reports proved to be strong predictors of subsequent recall. Left frontal activation has been associated with the organization of material to aid subsequent recall [43]. Introspection is particularly well suited to revealing strategic processes of this sort [44].
(4) Problem solving and planning [45]: numerous protocol analyses of problem-solving tasks have been conducted [11]. Goel writes that 'Most of our substantive results…stem from a combination of this methodology and information processing theory.' [46].
(5) Attentional processes [47]: introspective reports present the most direct method of assessing the 'foreground and background' of attention, and the temporal sequence of its 'flights and perchings', to use James's terms [1]. Lutz *et al*. [48] (and see below) show that detailed phenomenological reports can be used to assess level of preparation during a simple visual task. EEG (synchronization) and behavioural (reaction time) measures both correlated with subjective report.

Despite the relatively insignificant amount of time and effort currently put into the systematic collection of introspective evidence, it has proved itself to be informative about higher cognitive processes. There is every reason to believe that larger and more consistent efforts will be rewarded. Consequently, we propose a re-conceptualization of the standard 'cognitive mapping' approach outlined above.

From *stimulus–response* to *script–report*

In cognitive experiments, great emphasis is placed on the need to hold conditions constant. Thus, experimenters often take care to provide highly standardized instructions. Nonetheless, every researcher knows that subjects don't always follow these instructions. The script can easily be misunderstood, overruled by previous conceptions [49], altered or simply rejected [50]. None of this is necessarily visible in the observable behaviour. In addition, it is neither practicable nor desirable for the instructions to provide a complete description of the part the subject plays. We propose that in many experiments it would be advantageous to take a detailed retrospective report as the final part of the experiment. This provides an opportunity for the subjects to answer back, as it were; to allow them to describe the role they actually played and the experiences they had of playing it. This re-conceptualization involves an expansion of the standard stimulus–response paradigm to incorporate a larger frame: that of script and report (see Fig. 1). There are three advantages to this.

First, it provides additional evidence about the nature of the task, which is at least partially independent of the researcher's theoretical perspective. This is important because of the difficulties inherent in task analysis. Introspective reports offer a partially independent and supplementary source of information about the demands of particular tasks. The value of reports is especially obvious for experiments in which stimulus and response are held constant, but the task instructions are varied. It has been argued that this type of design is particularly valuable for imaging investigations of prefrontal function [36]. This design has been used in studies of attention [51], speech perception [52], memory [53], imagery [54], and theory-of-mind [49].

Second, it allows the experimenter to identify specific strategies used during task performance [11,44]

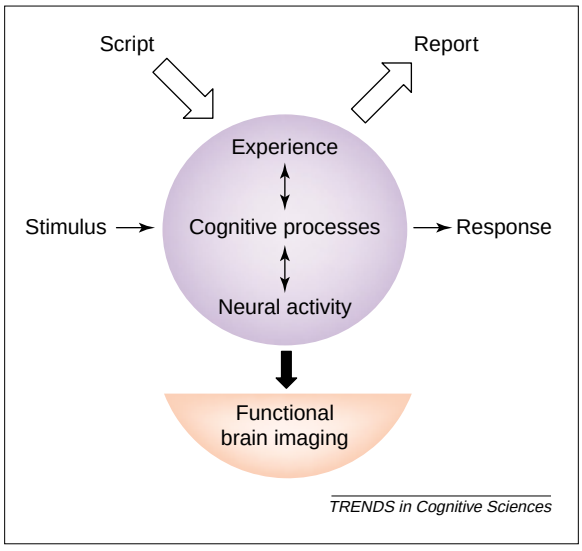

Fig. 1. Schematic illustrating a widening scientific perspective on cognitive processes. Post-behaviourist experimental psychology focused primarily on stimulus and response as a means of inferring cognitive processes. In recent years, there has been a strong shift towards the use of functional imaging evidence as a supplementary source of data to guide inferences about cognitive functions. We propose an even wider frame, explicitly acknowledging analysis of (the experimenter's) script and (the subject's) report as a further source of evidence about cognitive processes. We argue for more explicit discussion of the subject's experiences within experimental contexts. Acknowledging this larger frame of cognitive tasks might be particularly important for understanding higher cognitive functions, such as those associated with prefrontal cortex.

which can vary across subjects [49] and/or across experimental conditions [38]. Strategy differences are not necessarily reflected by report [55]. However, it is clear that important strategy differences are overlooked when reports are not elicited [38,42,56–59].

Third, and most important for the larger project alluded to here, this method serves to draw attention to the subject's experience. Post-experimental interviews (which clearly rely on episodic recall, hence our use of the term 'retrospection') are not well suited to probing the finer details of experience (such as variations between trials [48]), or to statistical analysis. However they can provide a 'crude' [60] or broad picture of the experiential states involved. At the very least, this should encourage experimenters to think more carefully before they speculate about the relevance of their objective measures to 'awareness' or other subjective phenomena. At best, retrospective reports will considerably enrich experimenters' understanding of 'what it is like' [61] to do the task, potentially revealing unexpected and important experiential phenomena. Experimenters can then frame detailed hypotheses regarding these phenomena, which can be tested using 'tailored' introspective methodologies.

Retrospective interviews do not interfere with the collection of 'objective' data, and so can be an uncomplicated add-on to virtually any experiment. The approach is illustrated by a recent brain-imaging investigation. Gallagher *et al.* [49] conducted a PET investigation of theory-of-mind where subjects played a computerized version of a simple game ('rock, paper, scissors'). In the critical experimental conditions, subjects were led to believe either that they were playing this game against another human, or that they were playing against a computer. Only one area, in para-cingulate cortex, showed differential activation in this condition, being more active when subjects thought they were playing a person. As the tasks involved identical stimuli and responses, some other evidence was needed to assess the psychological impact of this manipulation. We used retrospective reports collected through semi-structured interviews to help understand how subjects related to the different task contexts. The interview guide was generated from open-ended interviews conducted during pilot studies. This procedure helped to refine the task design, allowed us to identify strategy differences between subjects, and revealed that subjects experienced the conditions as clearly distinct.

The tailored use of introspective evidence involves asking subjects to make repeated reports in which they categorize or rate a specific aspect of their experience. The major advantage of such 'closed' questions is that the results are suitable for statistical analysis. However, there are also complications. First, introspecting is itself a cognitive activity and can influence other cognitive processes (see Box 2 and Schooler's article in this issue, pp. 339–344). Second, care must be taken that the categories or rating-scale used truly reflects the subject's experiences [62], and that it is used consistently across subjects. Subjects must still be interviewed, to check that the experimenter is not distorting the interpretation of the reports [63], or overlooking important phenomena.

Lutz *et al.* [48] provide an excellent example of the tailored use of introspective evidence, demonstrating that reports reflect variability in neuronal response even when the task is simple and does not change. Subjects were trained until they could recognize stable categories of experience, which related both to their state of 'preparedness' and their experience of the onset of the visual stimulus. Reports were then grouped post-hoc into 'phenomenological clusters' [64]. 'Well-prepared' clusters correlated with high gamma band EEG synchronization over frontal electrodes before stimulation, and with ERP and induced synchronization measures post stimulation.

Conclusion

We propose a rethinking of the standard cognitive mapping paradigm, which would render the mental processes studied in cognitive activation experiments subject to a methodological triangulation in which objective behavioural measurement, recordings of brain activity and introspective evidence can be related to each other. We provide evidence that higher

Acknowledgements
A.I.J. was supported by a grant from the Wellcome Trust throughout most of the preparation of this manuscript. The authors thank Chris Frith for his comments.

cognitive functions can be better understood when we broaden our perspective from its current narrow focus on stimulus–response to a wider frame, incorporating script–report. Clearly much more work needs to be done before we will understand the many ways in which these different layers of description fit together. However, we cannot see any logic to the current practice of ignoring one of these sources of evidence, the introspective report. As the biologist Seymour Kety noted, 'Nature is an elusive quarry, and it is foolhardy to pursue her with one eye closed and one foot hobbled' [65].

References

1 James, W. (1890) *The Principles of Psychology*, Holt, New York
2 Shallice, T. (1988) Information-processing models of consciousness: possibilities and problems. In *Consciousness in Contemporary Science* (Marcel, A.J. and Bisiach, E. eds), pp. 305–333, Clarendon Press
3 Frith, C.D. and Frith, U. (1999) Interacting minds – a biological basis. *Science* 286, 1692–1695
4 Praetorius, N. (2000) *Principles of Cognition, Language and Action: Essays on the Foundations of a Science of Psychology*, Kluwer Academic Publishers
5 Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press
6 Owen, A.M. (1997) The functional organization of working memory processes within human lateral frontal cortex: the contribution of functional neuroimaging. *Eur. J. Neurosci.* 9, 1329–1339
7 Elliott, R. *et al.* (2000) Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb. Cortex* 10, 308–317
8 Duncan, J. *et al.* (1996) Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* 30, 257–303
9 Hurlburt, R.T. and Heavey, C.L. (2001) Telling what we know: describing inner experience. *Trends Cogn. Sci.* 5, 400–403
10 Bisiach, E. (1988) The (haunted) brain and consciousness. In *Consciousness in Contemporary Science* (Marcel, A.J. and Bisiach, E., eds), pp.101–120, Clarendon Press
11 Ericsson, K.A. and Simon, H.A. (1993) *Protocol Analysis: Verbal Reports as Data* (*Revised edn*), MIT Press
12 Latour, B. (1993) *We Have Never Been Modern*, Harvester Wheatsheaf, New York
13 Varela, F.J. and Shear, J. (eds) (1999) The view from within. *Special issue of the Journal of Consciousness Studies* (Vol. 6, Issues 2–3), Imprint Academic
14 Nemirow, L. (1990) Physicalism and the cognitive role of acquaintance. In *Mind and Cognition: A Reader* (Lycan, W.C. ed.), pp. 490–499, Blackwell
15 Dixon, N.F. (1971) *Subliminal Perception: The Nature of a Controversy*, McGraw-Hill
16 White, P.A. (1988) Knowing more about what we can tell: 'introspective access' and causal report accuracy 10 years later. *Br. J. Psychol.* 79, 13–45
17 Lieberman, D.A. (1979) Behaviourism and the mind: a (limited) call for a return to introspection. *Am. Psychol.* 34, 319–333
18 Pickering, A. (ed.) (1992) *Science as Practice and Culture*, Chicago University Press
19 Holender, D. (1986) Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: a survey and appraisal. *Behav. Brain Sci.* 9, 1–66
20 Nisbett, R.E. and Wilson, T.D. (1977) Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 75, 522–536
21 Boring, E.G. (1953) A history of introspection. *Psychol. Bull.* 50, 169–189
22 Kellogg, R.T. (1982) When can we introspect accurately about mental processes? *Mem. Cogn.* 10, 141–144
23 Lyons, W. (1983) The transformation of introspection. *Br. J. Soc. Psychol.* 22, 327–342
24 Hurlburt, R.T. (1990) Sampling normal and schizophrenic inner experience. In *Emotions, Personality, and Psychotherapy* (Izard, C.E. and Singer, J.L. eds), Plenum Press
25 Jack, A.I. and Shallice, T. (2001) Introspective physicalism as an approach to the science of consciousness. *Cognition* 79, 161–196
26 Blumenthal, A.L. (1979) Wilhelm Wundt, the founding father we never knew. *Contemp. Psychol.* 24, 547–550
27 Harrison, J.E. and Baron-Cohen, S. (1997) Synaesthesia: an introduction. In *Synaesthesia* (Baron-Cohen, S. and Harrison, J.L. eds), pp. 3–16, Blackwell
28 Jack, A.I. (2001) Paradigm lost: a review of '*Consciousness Lost and Found*' by Lawrence Weiskrantz. *Mind Lang.* 16, 101–107
29 Kuhn, T.S. (1962) *The Structure of Scientific Revolutions* (3rd edn, 1996), University of Chicago Press
30 Frith, C.D. (2000) The role of dorsolateral prefrontal cortex in the selection of action, as revealed by functional imaging. In *Control of Cognitive Processes: Attention and Performance Vol. XVIII* (Monsell, S. and Driver, J., eds), pp.473–576, MIT Press
31 Duncan, J. and Owen, A.M. (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483
32 Pandya, D.N. and Yeterian, E.H. (1996) Comparison of prefrontal architecture and connections. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1423–1432
33 Burgess, P.W. (1997) Theory and methodology in executive function research. In *Methodology of Frontal and Executive Function* (Rabbitt, P. ed.), pp. 81–116, Psychology Press, Hove, UK
34 Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202
35 Shallice, T. and Burgess, P. (1996) The domain of supervisory processes and temporal organization of behaviour. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1405–1412
36 Fletcher, P.C. and Henson, R.N. (2001) Frontal lobes and human memory: insights from functional neuroimaging. *Brain* 124, 849–881
37 Vermersch, P. (1999) Introspection as practice. *J. Conscious. Stud.* 6, 17–42
38 Sperling, G. and Speelman, R.G. (1970) Acoustic similarity and auditory short-term memory: experiments and a model. In *Models of Human Memory* (Norman, D.A. ed.), Academic Press
39 Tulving, E. (1972) Episodic and semantic memory. In *Organization of Memory* (Tulving, E. and Donaldson, W. eds), Academic Press
40 Burgess, P.W. and Shallice, T. (1996) Confabulation and the control of recollection. *Memory* 4, 359–411
41 Kroll, N.E.A. and Kellicutt, M.H. (1972) Short-term recall as a function of covert rehearsal and of intervening task. *J. Verb. Learn. Verb. Behav.* 11, 196–204
42 Paivio, A. and Yuille, J.C. (1969) Changes in associative strategies and paired-associate learning over trials as a function of word imagery and type of learning set. *J. Exp. Psychol.* 79, 458–463
43 Henson, R.N. *et al.* (1999) Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *J. Neurosci.* 19, 3962–3972
44 Morris, P.E. (1981) Why Evans is wrong in criticizing introspective reports of subject strategies. *Br. J. Psychol.* 72, 465–468
45 Shallice, T. (1998) *From Neuropsychology to Mental Structure*, Cambridge University Press
46 Goel, V. *et al.* (1997) A study of the performance of patients with frontal lobe lesions in a financial planning task. *Brain* 120, 1805–1822
47 Posner, M.I. and Rothbart, M.K. (1998) Attention, self-regulation and consciousness. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1915–1927
48 Lutz, A. *et al.* (2002) Guiding the study of brain dynamics by using first-person data: synchrony patterns correlate with ongoing conscious states during a simple visual task. *Proc. Natl. Acad. Sci. U. S. A.* 99, 1586–1591
49 Gallagher, H. *et al.* Imaging the intentional stance. *Neuroimage* (in press)
50 Roepstorff, A. (2001) Brains in scanners: an Umwelt of cognitive neuroscience. *Semiotica* 134, 747–765
51 Corbetta, M. *et al.* (1991) Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J. Neurosci.* 11, 2383–2402
52 Zatorre, R.J. *et al.* (1992) Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849
53 Wagner, A.D. *et al.* (1998) Prefrontal cortex and recognition memory: functional-MRI evidence for context-dependent retrieval processes. *Brain* 121, 1985–2002
54 O'Craven, K.M. and Kanwisher, N. (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023
55 Siegler, R.S. and Stern, E. (1998) Conscious and unconscious strategy discoveries: a microgenetic analysis. *J. Exp. Psychol. Gen.* 127, 377–397
56 Siegler, R.S. (1987) The perils of averaging data over strategies: an example from children's addition. *J. Exp. Psychol. Gen.* 116, 250–264
57 Siegler, R.S. (1989) Hazards of mental chronometry: an example from children's subtraction. *J. Educ. Psychol.* 81, 497–506
58 Marquer, J. and Pereira, M. (1990) Verbal reports of cognitive strategies. *Q. J. Exp. Psychol.* 42A, 147–168
59 Morris, P.E. and Reid, R.L. (1970) Repeated use of mnemonic imagery. *Psychonomic Sci.* 20, 337–338

60 Schooler, J. (2002) Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends Cogn. Sci.* 6, 339–344
61 Nagel, T. (1974) What is it like to be a bat? *Philos. Rev.* 83, 435–450
62 Wilson, T.D. and Schooler, J.W. (1991) Thinking too much: introspection can reduce the quality of preferences and decisions. *J. Pers. Soc. Psychol.* 60, 181–192
63 Laming, D.R.J. (1997) *The Measurement of Sensation*, Oxford University Press
64 Varela, F.J. (1996) Neurophenomenology: a methodological remedy for the hard problem. *J. Conscious. Stud.* 3, 330–349
65 Kety, S. (1960) A biologist examines the mind and behavior. *Science* 132, 1861–1870

Re-representing consciousness: dissociations between experience and meta-consciousness

Jonathan W. Schooler

A distinction is drawn between non-conscious (unexperienced), conscious (experienced), and meta-conscious (re-represented) mental processes. There is evidence for two types of dissociations between consciousness and meta-consciousness, the latter being defined as the intermittent explicit re-representation of the contents of consciousness. Temporal dissociations occur when an individual, who previously lacked meta-consciousness about the contents of consciousness, directs meta-consciousness towards those contents; for example, catching one's mind wandering during reading. Once meta-consciousness is triggered, translation dissociations can occur if the re-representation process misrepresents the original experience, such as when one verbally reflects on non-verbal experiences or takes stock of subtle or ambiguous experiences.

Jonathan Schooler
518 Learning Research and Development Center, 3939 O'Hara Street, University of Pittsburgh, Pittsburgh PA 15260, USA.
e-mail: schooler+@pitt.edu

Philosophers and researchers have differentiated numerous forms of consciousness. Nevertheless, the conscious/non-conscious distinction, as determined by verbal reportability, represents the fundamental dimension upon which sentient experience is characterized. Critical to both the centrality of the conscious/non-conscious distinction, and its equation with reportability, is the assumption that people are explicitly aware of their conscious experiences. However, this assumption is challenged when subjective experience is dissociated from the explicit awareness of that experience. Such dissociations demonstrate the importance of distinguishing between consciousness and 'meta-consciousness' (the explicit awareness of the contents of consciousness), and illustrate the potential dangers of using self-reports as an index of consciousness.

Dissociations between consciousness and meta-consciousness

Although it generally seems that we are aware of the contents of experience, various situations illustrate dissociations between having an experience and knowing that one is having that experience. Such dissociations are exemplified by the situation of suddenly realizing that your mind has wandered while reading. Although the contents of such mind-wandering episodes are certainly experienced, the explicit awareness that your mind has wandered appears temporarily absent, as evidenced by the futility of continuing both activities. Other everyday situations similarly illustrate dissociations between experience and meta-consciousness, such as suddenly realizing that you have been successfully writing for hours, abruptly recognizing that you are extremely thirsty, or fiercely shouting '*I'm not angry!*'

The occurrence of experiences in the absence of their explicit recognition illustrates the value of distinguishing between consciousness and the explicit awareness of the contents of consciousness. Although the value of positing a meta-level of consciousness (also variously referred to as meta-awareness [1–3], reflective awareness [4,5], reflexivity [6], and reflexive consciousness [7]) has been noted in the past, more often, phenomenal experience (herein also referred to as consciousness) and meta-consciousness* are treated as one and the same. Moreover, when subtypes of consciousness are considered, meta-consciousness is typically ignored in favor of related constructs that differ in important respects (see Box 1). In the following discussion I outline a basic characterization of the relationship between non-conscious, conscious and meta-conscious mental processes (see Fig. 1) which, although rudimentary, suggests two potentially important ways in which consciousness and meta-consciousness can become dissociated.

The relationship between non-conscious, conscious, and meta-conscious processes

Conscious and non-conscious cognitive activities occur continuously throughout our waking hours. Much of the monitoring of consciousness is carried out by non-conscious processes that track goals [8], select strategies [9], and modulate the contents of thought [10]. Periodically attention is directed

*I use the terms meta-consciousness and meta-awareness interchangeably, as I find each term sounds better in certain contexts. In the future, a meaningful difference between the two terms might emerge, or one might prove preferable over the other.

http://tics.trends.com 1364-6613/02/$ – see front matter PII: S1364-6613(02)01949-6