mental disorder, requires a firm foundation of cognitive and clinical assessments. Posner and colleagues note that their clinical ‘...diagnosis involved 11 hours of objective measures and interviews...’ By contrast, assessment of the components of attention through administration of the ANT requires only 20 minutes or so [8]. Imagine how much more we will learn when exploratory measures like the ANT are followed up by a more thorough cognitive assessment as might be provided if 11 hours were spent on the cognitive side of the equation.

As illustrated in Fig. 2, a fruitful cognitive neuroscientific strategy [14] seeks to link mind and brain through computational models that explicitly reveal how cognitive processes are implemented in neural structures. A noteworthy feature of Posner’s research strategy is the diversity of sources of evidence that he and his colleagues have used to explore the mind–brain interface: neuroimaging, development, genetics, brain damage, psychiatric disorders and individual differences [8,9,11–13,15]. Through such converging evidence and with the development of computationally explicit theories (for example [16–18]) exploration will eventually give way to understanding.

References

How are qualia coupled to functions?

Jeffrey Gray
Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK

What determines the nature of subjective experience: associated behavioural functions or mediating neural activity? A recent analysis by Hurley and Noë of a variety of cases of behavioural and neural plasticity shows that, under different conditions, either can predominate. This adds to other efforts to transfer the doctrine of functionalism from philosophical debate to empirical scrutiny, where it is hoped that it may eventually be resolved.

No-one doubts that behavioural function is tightly linked to neural activity, nor is there any mystery in principle about how the linkage is accomplished (despite many details that remain to be established). But sometimes – although by no means always [1] – this cozy marriage turns into a troubling ménage-à-trois, joined by a mysterious intruder: conscious experience. Such experience, at least most of the time, if not invariably [2], takes a perceptual form. But there are also powerful modes of unconscious sensory processing [3], so the vocabulary of perceptual science is often ambiguous as to whether it addresses conscious or unconscious perceiving. For this reason, to refer to specifically conscious perceptual experiences, I shall borrow from philosophy (but without intending adherence to any particular philosophical position) the term quale (plural, qualia).

So, in this ménage-à-trois, do qualia couple principally with behavioural function or neural activity? The answer to this question given by the dominant view in contemporary science and philosophy – functionalism [4] – is behaviour. In its extreme form, this doctrine goes so far as to claim that the use of brain tissue to control behavioural function is incidental: if identical functions were accomplished by computers, robots or an array of tin cans, then they would give rise to identical states of consciousness. The contrary answer – that qualia are coupled to the
neural activity that underlies them – has yet to be even clearly articulated.

The case for functionalism

A recent article by two philosophers, Susan Hurley and Alva Noë [5], does the field sterling service by bringing these issues into the glare of empirical data-light. Drawing upon a variety of examples of neural plasticity, they distinguish between two ways in which qualia can be affected by changes in the normal relations between behaviour and neural activity: ‘cortical dominance’ and ‘cortical deference’.

In cortical dominance ‘cortical activation from a new peripheral input source gives rise to experience with a qualitative character normally or previously associated with cortical activity in that area’ (Ref. [5], p. 3). An example is a patient with a phantom limb [6] in whom, after amputation of an arm, stroking of the face is felt as the stroking of a phantom arm. In such cases it appears that activity in a region of the somatosensory cortex dedicated to sensations from the arm continues to produce the feeling of being stroked on the arm, despite the source of its activation (the face) now being different. So here qualia (feelings in an arm) are linked to neural activity, not to behaviour appropriate to peripheral input (stroking the face). Or, in Hurley and Noë’s terms, cortical neural activity is ‘dominant’ over functional input.

The converse case – cortical deference – is illustrated when congenitally blind persons read Braille. A number of studies (reviewed in Ref. [5]) show that such reading, felt by the blind person as tactile experience, is nonetheless mediated by activity in the visual cortex. So here tactile qualia are linked to the behaviour (touching Braille characters) and not to the neural activity that underlies them: activity in the visual cortex ‘defers’ to its unusual tactile inputs.

It is cortical dominance that gives functionalism trouble – cortical deference (qualia following function) is what this doctrine predicts. Hurley and Noë are themselves very much in the functionalist tradition. Not surprisingly, therefore, they see cortical deference as the norm, and dominance as ‘the exception that needs to be explained’ (Ref. [5], p. 29). Their account of cortical deference follows lines that each has explored before [7,8]. It proposes that qualia in general, and cortical deference in particular, depend upon dynamic sensorimotor interactions between the perceiver and the environment (a functionalist hypothesis par excellence).

In setting out this account, Hurley and Noë draw heavily upon examples in which plasticity is set in train, not by an abnormal brain (as in phantom limb or Braille reading), but by abnormal environmental contingencies, especially the now classical studies of adaptation to the wearing of various kinds of visually distorting prisms (reviewed in Ref. [7], Chapter 8). From these they conclude that cortical deference occurs only when the subject is able actively to explore the new sensorimotor contingencies that become operative after disturbance to the old or normal contingencies (whether the disturbance is due to alteration in the environment or rewiring of the brain). This summary is well justified in its application to the prism-adaptation studies. But it is less clear that it applies to all cases in which the disturbed contingencies depend upon abnormal brain wiring.

A case to which the ‘dynamic sensorimotor interaction’ account applies well is that of phantom limb. Because an amputated arm cannot be moved, there are no new sensorimotor contingencies to modify the old qualia (feelings in the arm) that stroking of the face now elicits (after new afferents from the face to the arm area of sensorimotor cortex have sprouted or been unmasked). But if you fake such new contingencies, as Ramachandran and Rogers-Ramachandran [6] did with a ‘mirror box’ (used to create an illusion of a seen intact hand in the position of the felt phantom hand), the phantom limb undergoes rapid modification. Here, then, cortical dominance, troubling for functionalism, is transmuted into unproblematic cortical deference; moreover, the transmutation fits well with an essentially functionalist theory.

The ‘problem’ of synaesthesia

This triumph for functionalist theory, however, is incomplete. The theory encounters a major obstacle in dealing with synaesthesia [9]. In ‘coloured hearing’, one of the most common forms of synaesthesia, subjects experience colours in response to heard words. This experience is accompanied by activity in the colour-selective visual area known as V4 or V5 [10], probably reflecting an aberrant, genetically determined [11], extra connection from cortical language areas directly to this region, V4/V5 (although this structural hypothesis has yet to be directly tested). We have here, then, a clear case of intermodal (audition to vision) cortical dominance. Now, a high proportion of coloured-hearing synaesthetes experience incongruent synaesthetic colours in response to colour names – for example, in response to hearing the word ‘red’ they have, say, a green colour experience – a phenomenon I have termed the ‘alien colour effect’ (ACE) [12]. One would expect this curious arrangement to give rise to cognitive interference. Sure enough, the speed of naming colours is reduced in proportion to the percentage of colour names affected by the ACE [13], thus behaviourally objectifying the synaesthetes’ self-reports, as in other cases of synaesthesia [9]. Further evidence supporting the self-reported ACE comes from unpublished experiments (D. Parslow et al., in preparation) in which we used fMRI to image the cortical activity of coloured-hearing synaesthetes while they performed a word–colour Stroop task [14]. Synaesthetes with the ACE showed a pattern of Stroop-specific cerebral activity, most notably in the hippocampus (consistent with a model of hippocampal mediation of the selection of the contents of consciousness [15]), that distinguished them, not only from non-synaesthete controls, but also from synaesthetes lacking the ACE.

Synaesthesia develops early in life: synaesthetes typically report that their unusual experiences go back as far as they can remember. Consider, then, Sally, a little girl with coloured-hearing synaesthesia and the ACE. Synaesthetes have normal colour vision, and they name colours and respond to such names appropriately [16]. So...
Sally will frequently have experiences that conform to the following sequence: ‘Look, Sally’, says her mother, ‘a red bus is coming round the corner’; because of her particular ACE, Sally has a green synaesthetic colour experience in response to the word ‘red’ and then, immediately afterwards, sees a red bus. What better set of ‘dynamic sensorimotor contingencies’ could there be to eliminate Sally’s disruptive green experience from the sequence? But it just doesn’t happen. Sally’s ACE remains with her for life, and green experiences continue to occur, automatically and unbidden (hallmarks of true conscious experience), whenever she hears the word ‘red’.

**Functionalism questioned**

Hurley and Noé, indeed, concede [Ref. [5], p. 30] that their model fails to account for synaesthesia. But they do not seem to realise how deep the problem goes. For synaesthesia runs counter to the most fundamental tenet of functionalism: that qualia are identical to the functions with which they are associated.

Suppose you ask the question: ‘what is the difference between experiences of the colour red and the colour green?’ Functionalism eliminates from the question the qualia of red and green as such. For these, it substitutes as explicandum the repertoire of behavioural responses by which the experiencing individual demonstrates the capacity to discriminate between red and green. This repertoire would include, for example, pointing to a red (or green) colour when requested to do so, using the words ‘red’ and ‘green’ appropriately in relation to the colours red and green, stopping at red traffic lights, stating that a lime is ‘green’ and a tomato, ‘red’, and so on.

Next, functionalism seeks an understanding of the mechanisms by which these behavioural ‘functions’ are discharged, using for example the familiar box-and-arrow diagrams of cognitive psychology, neural networks or computer simulations. A full ‘function’ for a given difference between qualia then consists of a detailed account of the corresponding differences in inputs, in outputs, and in the mechanisms that mediate between input and output. If such a full functional account is given, then according to functionalism, there is no further answer that can be given to the original question: what is the difference between the subjective experiences (the qualia) of red and green? To continue asking this question in the face of such a complete functionalist account would be a meaningless activity. For, according to functionalism, qualia just are the functions (input–mechanism–output) by which they are supported.

Given this conceptual framework, there should be no cases in which one quale is associated with two disparate functions. Yet that is just what happens in coloured-hearing synaesthesia. Colour qualia, affirmed by the synaesthete herself to be closely alike, occur in response to both heard words and seen coloured surfaces [13]. Nor, according to functionalism, should there be any cases in which two disparate qualia are associated with the same function. Yet, if we take seriously the Hurley–Noé hypothesis that qualia are determined by dynamic sensorimotor contingencies, that is again just what happens in coloured-hearing synaesthesia. Take Mary, a synaesthete in whom the spoken word ‘train’ elicits both the hearing of the word and a synaesthetic colour blue. There is only one set of dynamic sensorimotor contingencies that govern Mary’s interactions with vibrations in the air and the activation of receptors in the cochlea. Yet they give rise simultaneously to two very different qualia: the phonetics of ‘train’ and the colour blue.

Functionalism claims to provide a general account in principle for the coupling of qualia to brain-and-behaviour. Thus, even one clear negative instance is sufficient to topple it completely. Until recently, efforts to provide such an instance were purely philosophical. Transfer of these issues to the experimental laboratory is long overdue. Now that its time has come, it is gratifying to see the move being led by two philosophers.

**References**


1364-6613/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved. doi:10.1016/S1364-6613(03)00077-9