The Sound Of One Hand Clapping

Jeffrey Gray
Department of Psychology
Institute of Psychiatry
De Crespigny Park
London SE5 8AF
U.K.

j.gray@iop.kcl.ac.uk

PSYCHE, 8(11), October 2002
http://psyche.cs.monash.edu.au/v8/psyche-8-11-gray.html

KEYWORDS: Contents of consciousness, comparator model, familiarity, motor program, goal, rightness, tip of the tongue.

COMMENTARY ON: B. Mangan (2001). *Sensation's Ghost. The Non-Sensory "Fringe" of Consciousness*, PSYCHE, 7(18).

ABSTRACT: The 'non-sensory' feelings of familiarity, rightness and tip-of-the-tongue postulated in the target article all find a natural explanation within existing models, including Gray's (1995) comparator model, of the way in which top-down and bottom-up processing interact to select the contents of consciousness.

1. No Qualia -- No Problem

Much current research aims to increase the precision with which one can describe the three-way set of correlations that underlie the problem of consciousness: between functions, neural states and conscious experiences. Clearly, description of these correlations requires adequate description of each of the three dimensions that go into them. In this respect, the phenomenology of conscious experience has long lagged behind its two partners. By drawing attention to a relatively neglected aspect of conscious experience, Sensation's Ghost (SG) makes an important contribution to remedying this situation, and for that reason I warmly applaud.

However, no matter how sophisticated these correlations become, they will not by themselves answer the question: how does consciousness fit into the rest of the natural world? As always, we are unlikely to find an answer to this question until we have

properly understood the question itself. For this reason much of the intense discussion of the Hard Problem of consciousness over the last couple of decades has been directed to formulating it in a way that strips it down to its essentials. This is why there has been so much concentration on the nature of full-blown qualia: that is, what we see, hear, touch, taste etc when we interact consciously with the external world (or, properly speaking, with what our brains construct so well as a simulated external world that we take it to be real; Velmans, 2000).

If there were no qualia, there would be no Hard Problem of consciousness, indeed no problem of consciousness at all. There are no doubt many other problems about conscious experience, including the one that SG flags up. But it is a reasonable assumption that, if we once get a serious handle on how the brain creates qualia (I choose my words carefully, in full knowledge that many philosophers and some scientists will object strongly to them), these other problems will fall fairly easily into place. That is why there is such a concentration upon qualia - it is not simply a matter of research myopia. Now, of course, this assumption may be wrong. Mangan, indeed, makes exactly the opposite one: that it will be better to start with 'less obtrusive aspects of our phenomenology' (p. 1). Who knows? He may be right. It is only after you know the solution to a problem that you can be sure of the best way to approach it.

In the absence of qualia there would be, not only no problem of consciousness, but no mind/body problem either. We could then just get on with the accumulation of data to firm up the set of correlations between functions and brain states in the reasonable belief that this alone will provide all we need to know about the way the brain does the job of mind. For there is no conceptual difficulty in understanding how the brain can do the job of mind; and there are already a host of empirical demonstrations of how it actually does it (in neuroscience) or might do it (in psychology or computer science). Turning to the specifics of SG, among other things there is no problem in understanding in principle how the brain might compute non-conscious familiarity or rightness.

That there are 'feelings' of familiarity and rightness, in the sense that people are able to report them (which is still, but not I hope for ever, the best available empirical hall-mark of conscious experience) is beyond dispute. Also beyond dispute is that familiarity and rightness cannot be conceptually the same as one another. As Mangan points out, one can tell the difference readily enough between two equally familiar words, only one of which is right for the current gap in a stream of discourse; just as one can tell of two words, both with the right semantics for the gap, that one is more familiar than the other. Examples of this kind can be multiplied readily. But it does not follow from this conceptual difference that the brain uses radically different machinery to compute the two types of 'feeling'. Indeed, I shall try to show below that it uses essentially the same machinery for both.

I also find persuasive Mangan's argument that these 'feelings' are conscious, though they are not full-blown qualia. It is for this reason that I have put the scare quotes round 'feeling'. Whatever 'familiarity' and 'rightness' are, they are not like the full-blown feeling of being tickled, having an itch, or stroking velvet. What they seem to lack is the vivid qualities that are normally called 'sensory' (it is hard to capture, however, whether this is

a difference in kind or degree). So, let us by all means go along with Mangan in calling them 'non-sensory experiences'. But we then still need a terminology that distinguishes these experiences from qualia. The phrase 'non-sensory qualia', also used by Mangan, is for me a step too far in obscuring this distinction; for we should then always need, I suppose, to talk about 'full-blown qualia', or something like that, for what are at present qualia tout court. To avoid this, I shall in the rest of this commentary use 'NSE' as an acronym for 'non-sensory experiences'.

While I find Mangan's overall argument in favour of the concept of NSE persuasive, some of its aspects are troubling. Thus, there is an assertion (p. 5) that 'any experience that occurs in more than one sensory mode is non-sensory.' This is clearly false. As one among many possible examples, the well-known ventriloquist's illusion (one hears a voice coming from what one sees as moving lips, even though it is in fact coming from somewhere else) shows that the qualia that go to make up 'hearing someone speak' are a combination of experiences in two sensory modalities - but it is certainly not a 'non-sensory' experience. Another troubling aspect is the flirtation with the rightly discarded notion of sense-data or 'naked sensations' (p. 5). Thus Mangan writes (loc. cit.) that NSE can be thought of as 'the contents in consciousness which, when added to and merged with sensations, create perceptions.' Oddly, for an avowedly phenomenological approach, this is surely and simply phenomenologically wrong. Not that the phenomenology here is simple - it rarely is (one of the many reasons why, while one should never neglect phenomenology, the aim must be to transcend it with a theoretically viable and experimentally testable account of consciousness). My take on the phenomenology is itself complex and not easily summarised (I expand upon it in a forthcoming book). The following is a tenuously brief overview.

A first division is into those qualia that are endowed with meaning, that is, in the philosophical sense of the term (Searle, 1983) are 'intentional', and those that are not. This division corresponds, at least roughly, to the distinction between those contents of consciousness that make up the perceived external world, and those that make up the perceived inner body (the perceived external body is part of the external world). The brain, of course, constructs both the perceived external world and the perceived inner body, but it does so using largely different structures (Damasio, 2000) and largely different sensory channels (vision, audition, touch, olfaction and gustation for the former; visceral, nociceptive and feedback from the autonomic nervous system for the latter). So far as I can judge from my own phenomenology, contents of consciousness of the former type are almost invariably intentional, but those of the latter type usually not. (Compare the feeling of nausea, which just is that feeling, with seeing a car's number plate, which cannot be seen other than as a series of meaningful letters and numbers.) There are, to be sure, exceptions to these generalisations. The qualia that normally go to construct the external world can occur in almost total isolation from meaning. This is the case after certain kinds of brain pathology, producing the agnosias, and - more tellingly - for at least some normal forms of musical experience. It is much harder for me to identify cases of the reverse dissociation -- meaning without qualia. An anonymous reviewer of this commentary offers as an example the concept of the square root of -1. Sadly, I am too poor a mathematician to verify the assertion that this can be apprehended as a 'naked

concept', without any visualisation of the symbol i or square root signs. My personal phenomenology tells me that, in the case of the meaningful qualia that make up the external world, the relevant experience is always integrated: qualia come saturated with meaningful interpretation. It is not possible in experience to separate these from one another. But we are perhaps here up against the limits of unaided introspection.

It is of course possible conceptually to separate NSE, such as familiarity or rightness, from the qualia with which they are from time to time allied; and it is possible in principle to devise machinery that puts them together into the alliance. In the remainder of this commentary, I shall focus on a specific way in which this can be done.

2. The Comparator Model

The 'comparator' model I shall present is largely my own, but it shares with several other contemporary models features relevant to the position set forth in SG. Mangan states that his 'overall concern is to link non-sensory experiences to non-conscious processes that evaluate context information in the modern sense' (p. 6). The aim of the comparator model is to account for the selection of the contents of consciousness. Thus, for Mangan, there are three elements: (1) contextual non-conscious processes, (2) NSE and (3) qualia. The comparator model explicitly links the first and third of these. But, as I shall now show, it also provides a natural account of NSE, and it does this without addition or modification.

The comparator model was initially developed in order to account for a wide range of empirical data relating to the neuropsychology of anxiety (Gray, 1982a, b; Gray and McNaughton, 2000). Subsequently, it was employed in an account of the positive psychotic symptoms of acute schizophrenia (Gray et al., 1991; Gray, 1998), and this led in turn to an application to the selection of the contents of consciousness (Gray, 1995, plus commentaries and response).

The essential computational function discharged by the comparator is to compare, non-consciously and quite generally, information currently received via all thalamocortical sensory pathways (up to the level of neocortical analysis) with a prediction as to what that information should be. The prediction is based jointly upon previous stimulus-stimulus and response-stimulus regularities (stored as memories) under circumstances similar to those operating now; the circumstances 'operating now' are themselves defined by the output of the comparator at the preceding comparison process. In addition, the comparator takes account of the subject's ongoing motor program, as what the world will be like in the next moment depends upon what the subject is doing in this one. These processes occur on a time base of the order of 100 ms from the termination of one process of comparison to termination of the next. The output from the comparison process selects a series of items in the neocortical description of the sensory world in the light of their novelty/familiarity and predictedness/unpredictedness (these concepts are not identical to one another, in just the same way that Mangan's familiarity and rightness

are not identical concepts). The selection is biassed towards items which are novel, either because they occur despite not being expected or because they fail to occur despite being expected; and towards items which are goals or sub-goals for an ongoing motor program. The selected items are reactivated by feedback from the comparator system to those areas of the sensory neocortex (visual, auditory, somatosensory etc) in which they have just been non-consciously analysed. It is this reactivation by feedback from the comparator that selects these items for entry into consciousness. (I have no serious idea how such an 'entry into consciousness' actually occurs, but then neither does anyone else; this is the nub of the Hard Problem. Also, this brief account, based as it is upon my 1995 paper, needs supplementation to account for the fact that the items towards which conscious experience is biased - unpredicted events and predicted goals - fit into a much more extensive framework making up the constructed external world; I deal with this in a forthcoming book.)

The major justifications for this model, as applied to the problem of consciousness (it is justified in quite other ways in relation to anxiety and schizophrenia), are concerned with the 'lateness' of conscious experience. As thoroughly reviewed by Velmans (1991, 2000), conscious experience occurs too late (by one or several hundreds of milliseconds) to affect most on-line action or cognition - it occurs after the event. This poses a dilemma for those who, like myself, suppose that consciousness is a consequence of normal Darwinian selection. Such selection requires a survival function: the selected capacity must foster survival up to the time of reproduction and necessary parental care and/or these reproductive activities themselves. But most of the activities that plausibly provide such survival functions are just those which occur too rapidly for conscious awareness to affect them. The comparator model provides a potential solution simultaneously to the lateness of conscious experience and to the apparent lack of a survival function.

The model accounts for the lateness of conscious experience by the time it takes for the comparison process to result in reactivation of selected items of information in neocortical sensory systems so that these become qualia. Note that the time base of the comparator system formed part of the initial theory of the neuropsychology of anxiety (Gray, 1982 a,b), in which it was based upon quite different considerations. Thus, its application to the lateness of conscious experience is not arbitrary.

The model provides survival value to conscious experience by treating this as a 'late error detector'. Novelty and departure from prediction are key features biassing the selection of items by the comparator system for entry into consciousness. While this detection must necessarily occur after the event (novelty and departure from prediction must occur before they can be detected), it can provide the basis for re-evaluation of the motor program (or lack thereof) that led to the occurrence of error. These postulates, too, formed part of the original development of the comparator model as applied to anxiety.

Let us now return to Mangan's concern to 'link non-sensory experiences to non-conscious processes that evaluate context information in the modern sense' (p. 6). The comparator model does precisely this job: it evaluates context information, and does so in order to compute, in Mangan's terms, 'familiarity' (is this an item which, given the context, is

expected?) and 'rightness' (is this an item which, given the current motor program, is what should be happening now?). Whether 1982 (when the model was first proposed) counts as 'in the modern sense', I am not sure. But several more recent models travel essentially the same ground (see, for example, the summary article by Dehaene and Naccache, 2001, introducing a recent issue of the journal, Cognition, devoted to 'the cognitive neuroscience of consciousness'), the main difference lying in the postulated neural substrate of the systems that perform this type of computation. Where I stress the hippocampal system, more recent views tend to emphasise the prefrontal, anterior cingulate and/or parietal cortex. The precise anatomical localisation of the computations, however, does not bear upon the issues raised in SG. What does bear upon these issues is the emphasis in all these models upon the interaction of top-down (contextual) and bottom-up (perceptual) processing as giving rise to the contents of consciousness. But there is nothing new under the sun. There are clear antecedents to my 1982 formulation of this general idea, including for example Neisser's (1967) book, cited in SG, and Miller, Galanter and Pribram (1960).

We should also not lose sight of Jackendorff's important 1987 book. This gives a searching analysis of the level at which informational structure enters consciousness. Jackendorff proposes an 'intermediate-level' theory of consciousness. Roughly, this holds that the contents of consciousness reflect informational structures derived from a combination (within each perceptual modality) of bottom-up and top-down processing. Jackendorff argues that one is not normally (and perhaps never) aware of sensation unaffected by conceptual interpretation (cf my comment above on the snare and delusion of sense-data), nor of pure conceptual structure, but only of an admixture of the two that optimises the fit between them. This formulation is surely in general true. It is impossible to bring to conscious awareness (or so I find; see the discussion of the square root of -1, above) the pure conceptual structure represented by, say, the string 7 + 5 + 12 without this taking the form of either seeing or hearing the string, whether by means of externalised stimuli or 'in my head'. Mangan, in contrast, claims that there are exceptions to this rule. In particular, he claims, the pure conceptual structures that can be expressed in English as 'this is familiar', 'this is unfamiliar', 'this is right' or 'this is wrong' can be consciously experienced in the absence of qualia: as 'naked concepts', one might say.

The comparator model, then, provides machinery to produce computations of familiarity and rightness, and it explains the relationship between this machinery and the selection of the full-blown qualia. Where, in this scheme of things, can we situate NSE? As we try to do so, let us respect Occam's razor: we should avoid, if possible, postulating either new machinery or new states of consciousness.

In fact, there is a very natural way to slot the NSE of familiarity and rightness into the comparator model. The tip of the tongue (TOT) phenomenon, discussed in SG, provides an excellent illustration of the slot. The hypothesis is that: (1) this phenomenon consists in the top-down computation of the predictive outputs of the comparator system; while (2) these are in the process of being fed into the perceptual systems (in the instance, those concerned with auditory language processing); but which (3) have not yet encountered

informational structures that fit the prediction; and which (4), when they are encountered, will become qualia (that mysterious 'entry into consciousness again!).

Now, to make this hypothesis clear, there is one thing to add to the outline of the comparator model above. This model was developed to account for what happens when perceptions are formed as the result of external stimulation of sensory surfaces. The reason to start here, as noted above, is that this is the kernel of the Hard Problem. However, there are of course plenty of conscious experiences that start with internal processing and are not supported by external stimulation. Let me use the terms 'public space' and 'private space' qualia to distinguish these two cases. (This distinction, by the way, is independent of the one between the perceived external world and inner body made earlier in this commentary.) Mental images, thoughts, recalled memories are all examples of private space qualia. I see no reason to suppose that the basic machinery for the selection of the contents of consciousness should differ between the private and public spaces. But there must necessarily be a different dynamic between the contributions of top-down and bottom-up processing in the two spaces. For, in the case of private space qualia, top-down processing, in addition to its central function of producing a predictive template against which to match incoming sensory information, must now use that template also to seek out and activate the very sensory informational structures that need so to be matched. (The current consensus is that these processes are largely carried out in the prefrontal, cingulate and parietal cortices, to which I add the hippocampal system; see Dehaene and Naccache, 2001.) This is no doubt why mental images are never (in the absence of pathology, e.g., hallucinations) as vivid as percepts based upon actual sensory input. This subjective difference finds echo in recent functional neuroimaging studies in which the brain activation patterns in sensory systems differ between perceived and imagined stimuli (e.g., Nunn et al., 2002, and references therein). Presumably, a systematic difference between the neural and psychological reactions to real and imagined stimuli, respectively, is essential to the brain's ability to construct a perceptual model of the real world and to keep this separate from other forms of perceptual construct which have a more indirect relationship to that world.

This gloss apart, the application of the comparator model to the TOT is, I think, straightforward. The TOT reflects the active formation of the predictive template and its persistence during an interval prior to its being successfully matched against an informational structure in perceptual systems. It lacks qualia until a successful match is made. On this hypothesis, therefore, there is no need to postulate any additional machinery, over and above that already included in the comparator system.

It still remains to ask, however, whether we should follow Jackendorff and say that this 'naked concept' (familiarity, etc) is non-conscious, or follow Mangan in calling it a non-sensory experience. In the attempt to resolve this remaining issue it is worth looking to certain phenomena in the pathology of vision as a possibly instructive parallel. In presenting this parallel, I follow closely a discussion by Dehaene and Naccache (2001, pp. 16-19). It is well known in clinical neuropsychology that the same general type of impairment may in some patients be accompanied by knowledge of the impairment, whereas others are ignorant of it. The latter condition is termed 'anosognosia'. In this

context, Dehaene and Naccache consider two types of blindness for part of the visual field. In one, scotoma due to damage to part of the retina, the patient is aware of his blindness in the corresponding part of the visual field. In the second, visual neglect after lesions to the parietal cortex, the patient is blind in regions of space contralateral to the site of the lesion, but is unaware of this fact. Applying the analogy back to the TOT, one may say that, in this, there is a (temporary) 'internal deafness' - there is a word one is trying to 'hear' inside one's head but one cannot (yet) hear it. There is also an awareness of this deafness; this awareness is the tip-of-the-tongue phenomenon itself. So, the case in the Dehaene and Naccache comparison which parallels TOT is the retinal scotoma.

Here, then, is how Dehaene and Naccache explain the difference between blindness with and without anosognosia. Their account of visual neglect treats the parietal cortex as part of the top-down processing involved in visual perception, responsible, more precisely, for an 'attentional amplification' of visual information that is required for entry into the 'global workspace' (a currently fashionable euphemism for 'consciousness'). Damage to this region would be expected, therefore, to give rise to two effects: loss of conscious visual percepts due to the lack of attentional amplification, and lack of attempts to induce such attentional amplification. The latter implies, in line with clinical observations, that the patient would be unaware of the fact that he is blind in the affected field. In the case of a retinal scotoma, in contrast, all top-down processing remains intact, but there is no bottom-up activity with which this processing might interact. It is the intact top-down processing which provides the patient with knowledge of his blindness. In the TOT, which of course takes place in people with intact brains, we may similarly assume that top-down processing is intact, but there is a temporary absence of bottom-up processing to meet it. The intact top-down processing provides the subject with knowledge that there is a gap waiting to be filled by the absent bottom-up processing.

To take the analogy one step further, might there be cases that are to the TOT what visual neglect is to a retinal scotoma? A case of this nature would be a person unable to find the right word but unaware of this impairment. Clearly, such cases exist in unfortunate abundance, though so far as I know only after damage to the brain, as for example in jargon aphasia.

We seem, then, to have a plausible way of explaining the TOT (and by extension, other NSEs discussed in SG, such as the feelings of familiarity and rightness) without going beyond concepts and processes already shown to do useful work in relationship to consciousness. And Occam's razor threatens us if we try to add any surplus. Compared to this satisfactory outcome, it is of minor importance whether one wishes to accept or not Mangan's description of these phenomena as 'non-sensory experience'. Within the comparator model, they represent outputs from a non-conscious system (the comparator itself) while it is participating in (familiarity or rightness combined with matching sensory input) or is about to participate in (the TOT) a process by which a set of qualia are constructed and selected for entry into consciousness. Note that, viewed in this way, NSE cut across the distinction drawn above between the public and private spaces of consciousness. Feelings of familiarity or rightness can accompany qualia that originate in either public space (I see a familiar face in the crowd) or private space (I hear a familiar

tune in my head), and in both cases one may perhaps be able to identify the origin of the familiarity (it's 'a chap from work' or 'a tune from La Traviata') or not. In the TOT, the NSE occurs in the absence (yet) of any bottom-up counterpart. I am reminded of the Zen 'sound of one hand clapping'. I cannot think of any better way of describing that sound than as 'a non-sensory experience' - perhaps this is as good a guide as any to the 'rightness' of Mangan's choice of phrase.

3. Coda: Psychopathology

I finish with some brief comments on psychopathology. Mangan relates NSEs to both anxiety (p. 21) and schizophrenia (p. 20). His comments in both instances are in general agreement with the way in which the comparator model has been applied to these two conditions.

As noted above, the model was first developed as part of an overall theory of the neuropsychology of anxiety. Within that theory (Gray, 1982a, b; Gray and McNaughton, 2000), generalised anxiety (equivalent to Mangan's 'free-floating anxiety') is treated as reflecting a state of chronic overactivity in the comparator, biassed towards monitoring the environment for potential sources of threat. In the application of the model to the positive symptoms of acute schizophrenia (Gray et al., 1991), these are treated as reflecting functional disconnection between that part of the comparator system which computes past regularities of experience and that part which conducts the comparison process. In consequence, the patient misconstrues stimuli or events, which for normal people would be familiar/predicted, as being novel/unpredicted. This hypothesis has been subjected to extensive empirical research (Gray, 1998) and is able to account for a wide range of positive schizophrenic symptoms (Hemsley, 1996). In its application to both conditions, detailed proposals have been made as to the neural systems that discharge the functions of the comparator model and as to the their pathology.

References

Damasio, A. (2000) *The feeling of what happens.* London: Heinemann.

Dehaene, S. & Naccache, L. (2001). Towards a cognitive science of consciousness: basic evidence and a workspace framework. *Cognition, 79,* 1-37.

Hemsley, D. R. (1996). Schizophrenia: a cognitive model and its implications for psychological intervention. *Behaviour Modification, 20,* 139-169.

Gray, J. A (1982a). *The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system.* Oxford: Oxford University Press.

Gray, J. A. (1982b). Précis of 'The Neuropsychology of Anxiety: an enquiry into the functions of the septo-hippocampal system'. *Behavioral and Brain Sciences, 5,* 469-484.

Gray, J. A. (1995). The contents of consciousness: a neuropsychological conjecture. *Behavioral and Brain Sciences, 18,* 659-676.

Gray, J. A. (1998). Integrating schizophrenia. *Schizophrenia Bulletin, 24,* 249-266.

Gray, J. A, Feldon, J, Rawlins, J. N. P., Hemsley, D. R. & Smith, A. D. (1991). The neuropsychology of schizophrenia. *Behavioral and Brain Sciences, 14,* 1-20.

Gray, J. A. & McNaughton, N. (2000). *The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system.* (2nd ed.) Oxford: Oxford University Press. 2000.

Jackendorff, R. (1987). *Consciousness and the computational mind.* Cambridge, Mass.: MIT Press.

Miller, G. A., Galanter, E. H. & Pribram, K. H. (1960). *Plans and the structure of behavior.* New York: Rinehart and Winston.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., Morris, R. G., Bullmore, E. T., Baron-Cohen, S., Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature Neuroscience, 5,* 371-375.

Searle, J. R. (1983). *Intentionality.* Cambridge: Cambridge University Press.

Velmans, M. (1991). Is human information processing conscious? *Behavioral and Brain Sciences, 14,* 651-669.

Velmans, M. (2000). *Understanding consciousness.* London: Routledge.