Neural response suppression, haemodynamic repetition effects, and behavioural priming

R.N.A. Henson a, b, *, M.D. Rugg a

a Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK
b Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College, London, UK

Abstract
Repeated stimulus processing is often associated with a reduction in neural activity, as measured by single-cell recording or by haemodynamic imaging techniques like PET and fMRI. These reductions are sometimes linked to the behavioural phenomenon of priming. In this article, we discuss issues relevant to theories that attempt to relate these phenomena, concentrating in particular on the interpretative limitations of current imaging techniques.

© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Faces; Fusiform; fMRI; ERP; MEG; Implicit

1. Introduction
A common finding in PET and fMRI studies using indirect (or “implicit”) memory tasks is a reduced haemodynamic response for repeated relative to unrepeated stimuli. This “repetition suppression” has been associated with the behavioural phenomenon of priming, as typically indexed by shortened reaction times, or improved identification of degraded stimuli [33]. A potential neural analogue of the haemodynamic decrease is “response suppression” [7], which refers to a reduced firing rate of neurons on stimulus repetition, as typically recorded in inferior temporal regions of the nonhuman primate (Fig. 1A).

Such parallels offer the attractive possibility of developing theories that bridge these different levels of neuroscience. Wiggs and Martin [39], for example, extending ideas of Desimone [7], proposed that repeated processing of a stimulus produces a “sharpening” of its cortical representation, whereby neurons coding features unnecessary for processing that stimulus respond less (i.e. exhibit response suppression). This results in a decrease in the mean firing rate of a population of neurons, and hence a decrease in the haemodynamic response from that region of cortex (Fig. 1B). Furthermore, they proposed that this "sparser" representation allows more efficient stimulus processing, which can result in behavioural priming effects.

While we are very much in favour of such theoretical efforts, the purpose of the present article is to raise some cautionary issues, particularly concerning the interpretation of neuroimaging data. We concentrate on the effects of face repetition on the haemodynamic response of a right fusiform region. This so-called “Fusiform Face Area” has been well-studied [22], though its precise functions are still debated [11,16,25]. Nonetheless, we believe our arguments extend to the more general goal of relating neurophysiological, neuroimaging and psychological data.

2. Mapping between neural firing and haemodynamic response
One set of issues concerns the mapping between neural firing and the haemodynamic response. Foremost, it must be remembered that, even with event-related fMRI, the haemodynamic response represents the integration of several seconds of neural/synaptic activity. This means that a decrease in the magnitude of a haemodynamic response may reflect a shortened duration of neural/synaptic activity, in the absence of a change in instantaneous firing rates (Fig. 2A).

* Corresponding author. Tel.: +44-20-7679-1131/7833-7483; fax: +44-20-7813-1420.
E-mail address: csonen@ucl.ac.uk (R.N.A. Henson).
Fig. 1. (A) An example of response suppression (or “decremental responses”) recorded from a monkey perirhinal neuron to first (top) and repeated (bottom) presentations of pictures during a serial recognition task (reprinted with permission of Elsevier Science from Brown and Xiang [4]). These decreases are typically seen in a relatively large number of neurons (e.g. approximately 25% of neurons recorded, or 50% of those visually-responsive) in many inferotemporal regions. The decreased firing rate is not nonspecific habituation, because it occurs after a single exposure to a stimulus and does not affect the firing rate to other stimuli for which the neuron is responsive. Moreover, it can last hours/days and numerous intervening stimuli (depending on the brain region, see text). (B) Schematic of the “sharpening” theory of Wiggs and Martin [39], showing the effect of repeating an unfamiliar face (left column) on the number of responsive neurons (middle column; neurons showing response suppression are greyed out) and the magnitude of the expected haemodynamic response (right column).
For example, Henson and Rugg [20] found that repetition of famous faces decreased not only the peak magnitude of the BOLD response in a right fusiform region (repetition suppression), but also its peak latency (Fig. 2B). The most parsimonious account of this combined change is that repetition caused a reduction in the duration of evoked neural/synaptic activity (see Fig. 2A legend). This interpretation is consistent with models that attribute priming to reduced “settling” times in attractor neural networks, owing to synaptic changes following processing of the prime [3]. These reduced settling times occur with a decrease in mean firing rate (or increased sparsity of representation) proposed by Wiggs and Martin.

A second consequence of its integrative nature is that the haemodynamic response is insensitive to whether repetition effects are early (e.g. “first pass”) or late (e.g. “reentrant”) to synaptic changes following processing of the prime [3]. These reduced settling times occurred approximately 250–300 ms. Similar findings were reported by Dale et al. [6], who used fMRI data to confirm that the source of extracranial MEG priming-related effects from human inferior temporal cortex and found face-specific potentials onsetting 150–200 ms post-stimulus, but little evidence that these potentials were modulated by repetition or familiarity (Fig. 2C). The earliest priming-related modulation (for famous faces preceded by a matching name) emerged approximately 250–300 ms. Similar findings were reported by Puce et al. [28], who used fMRI data to constrain the source of extracranial MEG priming-related effects during semantic decisions on words. They found an initial “wave” of activity that had spread to temporal, parietal and frontal regions by 185 ms (Fig. 2D). The earliest repetition effect however emerged in a left anterior inferior temporal region at 250 ms, and did not peak in most regions until approximately 400 ms. Data like these raise the possibility that repetition-related “deactivations” in some regions (e.g. medial temporal lobes), for example, have been interpreted as automatic novelty-related “activations” to initial presentations of stimuli, rather than priming per se [15].

A second, related issue concerns whether repetition suppression is the cause or effect of priming. Haemodynamic decreases associated with perceptual priming, for example, might simply be a consequence of reduced gaze duration, and associated priming, should be greater for repeated unfamiliar stimuli than for repeated familiar stimuli. Henson et al. [17] however found the opposite pattern in the face-evoked BOLD response of a right fusiform region (a likely candidate for the formation of perceptual representations of faces [12]). This region showed greater repetition suppression for famous faces than unfamiliar faces (Fig. 3B). Though this pattern may have other explanations (e.g. relating to the specific type of face representation, or top–down feedback from more anterior temporal regions), it is consistent with alternative “abstractionist” theories of priming [36]. According to these theories, priming reflects lowered thresholds for activating pre-existing representations in memory. Thus, only familiar faces, with pre-existing representations, can be primed, consistent with some behavioural data [9]; though see [13]).

3. Mapping between haemodynamic response and behaviour

Priming is claimed to reflect the operation of unconscious or “implicit” memory [34]. This claim is based in part on the fact that an indirect memory task was used to elicit the effect. Though indirect tasks do not necessitate detection of stimuli, the behavioural and imaging data obtained in such tasks may represent a contribution from voluntary or involuntary explicit memory [30]. Only a few PET and fMRI studies have tried to control for such explicit memory “contamination” by using, for example, subliminal presentation of the prime [27], or an experimental manipulation known to affect implicit and explicit memory differentially [1]. The association of priming with repetition suppression would benefit from further imaging studies that, for example, employ amnesic patients [5], or use pharmacological manipulations thought to selectively affect explicit memory [37]. Moreover, not all brain regions showing repetition suppression necessarily reflect the same neurophysiological mechanism, or have the same behavioural consequences; repetition-related “deactivations” in some regions (e.g. medial temporal lobes), for example, have been interpreted as automatic novelty-related “activations” to initial presentations of stimuli, rather than priming per se [15].

A second, related issue concerns whether repetition suppression decreases associated with perceptual priming, for example, might simply be a consequence of reduced gaze duration,
or diminished attention, to repeated stimuli. In other words, such changes might be an effect rather than cause of the behavioural priming effect (the cause presumably arising elsewhere in the brain). In response to this argument, it is often noted [33] that repetition suppression is rarely observed in early sensory regions, where one might also expect attentional effects, or in motor regions, where one might expect effects of priming-related decreases in reaction time. Nonetheless, the question of cause and effect cannot be fully resolved with correlational techniques such as fMRI and PET, and the association of priming with repetition suppression would benefit from supplementary evidence that a region showing repetition suppression also produces a priming deficit when lesioned [10] or targeted by TMS.

A third issue relates to the automaticity of repetition suppression. Wiggs and Martin state that repetition suppression "...happens automatically in the cortex..." and is "...an intrinsic property of cortical neurons providing a form of perceptual learning that allows us to identify previously encountered objects quickly" ([39], p. 231). If so, one might expect to see repetition suppression in perceptual regions whenever a stimulus is perceived repeatedly, regardless of the cognitive task performed on that stimulus. However, several imaging studies have shown that repetition-related responses in perceptual regions can vary with the task, particularly whether it is a direct or indirect memory task (see [19] for a review). Henson et al. [18], for example, found that the repetition suppression (for famous faces) that was seen in a right fusiform region in an indirect task (fame-judgement) was not seen in a direct task (episodic recognition, Fig. 3C). While there may be several explanations of this particular task-by-repetition interaction (such as the "swamping" of automatic suppression by additional top-down activation in the direct memory task), the task-sensitivity of haemodynamic repetition suppression cautions against the assumption that it is an obligatory consequence of stimulus repetition (an assumption often implicit in the use of repetition suppression as a tool to identify separate stages in, for example, visual object processing, e.g. [14]).

Finally, a small number of imaging studies have found haemodynamic increases, rather than decreases, in the context of repetition priming. Dolan et al. [8], for example, found haemodynamic increases in the right fusiform for repeated versus initial presentations of degraded images of unfamiliar faces when undegraded versions intervened (Fig. 3D). This "repetition enhancement" indicates that haemodynamic decreases are not associated with all forms of repetition priming (see [19] for further discussion).

4. Mapping between neural activity and behaviour

A final set of (more tentative) issues relate to the mapping between neural firing and behaviour. Foremost, the association of priming with neural response suppression, as recorded in monkey cortex, is yet to established, not least because "priming" is rarely studied in animals. Indeed, the same phenomenon of response suppression in monkey anterior temporal cortex is interpreted in terms of explicit, recognition memory [4].

Secondly, Wiggs and Martin [39] do not offer a mechanistic account as to why sparser representations necessarily allow faster/more efficient processing of stimuli. How a sparse (or distributed) representation affects the speed with which a neural network can identify a stimulus would appear to depend on the dynamics of the network. Synaptic changes within an attractor network, for example, may produce faster relaxation times without a change in the sparsity of firing (see earlier).

Thirdly, the duration of response suppression in monkey inferior temporal cortex typically decreases from anterior regions (e.g. 24 h in perirhinal cortex, 10 min in area TE, [4]) to posterior occipitotemporal regions (where it may not survive a single intervening stimulus [2]). However, many perceptual priming effects in humans have been attributed...
but not a direct repetition–detection task (right) in the same right fusiform region as (B), from Henson et al. [18]. (C) BOLD repetition suppression for familiar (but not unfamiliar) faces during an indirect fame–judgment task (left) afterwards (Fa), in right fusiform region shown in transverse section (reprinted with permission of MacMillan Magazines Limited from Dolan et al. [8]).

References

5. Conclusions
We have presented several issues relevant to theories of priming that attempt to link neural, haemodynamic and behavioural levels of neuroscience. Some of these issues are problematic for theories like that of Wiggs and Martin [39]. Though none of the issues refutes the theory, they suggest that the theory is insufficient to account for the range of priming-related neural effects. Though we cannot presently offer an alternative theory that addresses the issues completely, we propose that such a theory would emphasise how repeated stimulus processing alters the strength of synap- tic transmission between different regions, which might af- fect the dynamics of a network settling on an interpretation (e.g. recognition) of that stimulus, and hence influence beh- avioural responses such as reaction times.

Finally, we note that we are not against the Wiggs and Martin theory in principle. We have discussed it here because it is to some extent typifies the ideas held implicitly by many researchers using functional neuroimaging. Indeed, we be- lieve Wiggs and Martin have greatly benefited the field by making these ideas explicit. As such, the theory has proved a useful vehicle for stimulating our thoughts and for raising more general issues concerning the goal of relating neuro- physiological, neuroimaging and psychological data.

Acknowledgements
This work is funded by Wellcome Trust Fellowship 060924. RH and MR are supported by the Wellcome Trust.

Fig. 3. (A) Schematic of apparent prediction of Wiggs and Martin [39], that the difference between successive repetitions decreases as stimuli become familiarised through multiple repetitions (see also [23]). (B) Repetition suppression for feet vs. second presentations of familiar faces (Ft) and F2, top graph, jagged lines show data binned every 2 s, smooth lines show linear model fit) but not unfamiliar faces (U1 and U2, bottom graph) during an indirect monitoring task in a right fusiform region (Talairach coordinates ±45 –37 –24) indicated by crosshair on inferior view of brain with cerebellum artificially removed, from Henson et al. [17]. (C) BOLD repetition suppression for familiar (but not unfamiliar) faces during an indirect fame–judgment task (left) but not a direct repetition–detection task (right) in the same right fusiform region as (B), from Henson et al. [18]; data binned every 1 s. (D) PET adjusted response to a degraded (binarized) image of an unfamiliar face, before presentation of intact version (Fb), and to a repeated presentation of that image afterwards (Fe), in right fusiform region shown in transverse section (reprinted with permission of MacMillan Magazines Limited from Dolan et al. [8]).
Henson RN, Shallice T, Gorno-Tempini ML, Dolan RJ. Face repetition effects in implicit and explicit memory tests as measured by fMRI. Cerebral Cortex 2002;12:176–86.


