1. Introduction

Most functional magnetic resonance imaging (fMRI) studies attempt to identify the brain regions involved in performing a task; that is, these studies are concerned with the locus or loci of on-line processes. The successes and limitations of this research programme have been discussed and debated at length (e.g., Henson, 2005; Page, 2006; Coltheart, 2006; Logothetis, 2008; Vul and Kanwisher, 2010). Over the last few years there has been an explosion of fMRI research designed to identify the brain regions involved in learning visual (e.g., Schwartz et al., 2002; Maertens and Pollman, 2005; Yotsumoto et al., 2008), auditory (e.g., Jäncke et al., 2001), motor (e.g., Parsons et al., 2005; Graydon et al., 2005), language (e.g., Mochizuki-Kawai et al., 2006; Wong et al., 2007), and other skills. These studies attempt to determine where on-line neural processes leave their mark in long-term memory, such that performance is improved following training. Here we consider this latter research programme, and argue that there are fundamental methodological and conceptual problems with these studies that render their conclusions regarding the locus of learning problematic.

2. Blood-oxygenation-level-dependent (BOLD) signals and their relation to neural changes that mediate learning

A key assumption motivating fMRI studies of on-line processing is that there is a close link between BOLD responses and the neural processes that support task performance. Indeed, this link enables fMRI to locate the brain areas that support on-line processing. In the same way, when using fMRI to study the locus of learning (or equivalently, the locus of neural plasticity), there must be a close link between the BOLD signal and the underlying neural processes that mediate learning. This assumption underpins research in all domains of learning, but below we focus on perceptual learning that is defined as practice-induced improvement in the ability to perform specific perceptual tasks (Ahissar and Hochstein, 2004) and is mediated by various modifications in the brain, from molecular changes in gene expression, cellular changes in numbers of synapses or dendritic length, to changes in organization of cortical maps (Buonomano and Merzenich, 1998; Kolb and Whishaw, 1998; Redondo and Morris, 2011). If fMRI is to identify the brain regions that support perceptual learning, then the BOLD signal needs to be sensitive to these physiological changes.

The most straightforward way to satisfy this requirement would be that the BOLD signal is driven by the metabolic costs of learning, e.g., the energy requirements associated with gene expression, the modifications of synapses, etc. However, there is no evidence that this is the case. Indeed, the neuronal modifications that mediate learning can take hours, days or even months following training, and as a consequence, studies relying on BOLD signal are not able to measure the costs of neural plasticity per se (Poldrack, 2000; Kelly and Garavan, 2005). Accordingly, researchers have adopted a more indirect method of linking BOLD responses to the locus of learning. That is, the BOLD signal is used to measure the locus of on-line processes at time 1 and 2, and any changes in BOLD levels from time 1 to time 2 are taken to reflect the locus of learning. For example, if area V1 produces weaker (or stronger) BOLD signals in a perceptual discrimination task at time 1 compared to time 2, then this change is taken to reflect learning within V1 (e.g., Furmanski et al., 2004; Maertens and Pollman, 2005).

Our key claim is that this line of inference is unsafe given that learning in one brain area (e.g., V1) may impact on-line...
processing across a wide range of areas (e.g., V2). Indeed, one key insight from fMRI research is that many areas of the brain are activated in the completion of any task, with evidence for both bottom–up and top–down on-line feedback (e.g., Wright et al., 2008; Beck and Kastner, 2009). Accordingly, any changes in the BOLD signal in V1 from time 1 to time 2 may reflect neural plasticity (e.g., molecular, synaptic and dendritic changes) within V1, or alternatively, may reflect plasticity somewhere else which nevertheless impacts on the V1 BOLD signal due to on-line feedback. That is, the relationship between the locus of learning and the BOLD signal change may be quite weak (see Fig. 1). We flesh out this argument by considering two specific studies of visual perceptual learning, but would like to emphasize that our analyses apply to all studies that associate BOLD signals to the locus of learning (perceptual or otherwise).

3. Prototypical findings and flaws in their interpretation

Let us briefly review two fMRI studies that report three common patterns of BOLD changes associated with visual learning, namely: (1) changes in BOLD activation across low-level and high-level visual areas, (2) BOLD changes restricted to high-level visual areas, and (3) BOLD changes restricted to low-level visual areas.

A study by Kourtzi et al. (2005) investigated the locus of learning new shapes in low-salience and high-salience conditions when shapes were embedded in cluttered scenes. Changes in BOLD signal for trained and untrained stimuli were recorded before and after 3 days of training. The low-salience condition was associated with increases in signal both in retinotopic cortex (including V1, V2, Vp, and V4v) and lateral occipital complex (LOC), whereas the high-salience condition was associated with selective decreases of signal in LOC. Based on these findings it was concluded that training in the low-salience condition resulted in learning in low-level and high-level visual areas (with learning leading to greater overall neural firing, with new cells recruited to the task), whereas learning in the high-salience training condition was restricted to LOC (with learning leading to less overall neural firing, due to the development of sparser codes).

However, these findings do not allow us to distinguish between several possible scenarios regarding the relationship between BOLD signal and the locus of learning, as depicted in Fig. 1. In the low-contrast condition, it is not clear whether the increased activation in a particular brain region (e.g., V2) was due to learning in that region, or whether it reflects interactions with other areas in which learning has taken place. For instance, learning might have been circumscribed to V1, with activation changes in other areas (such as V2) reflecting knock-on effects [as in scenario (c)].

![Diagram](image-url)
Similarly, the results in the high-contrast condition are ambiguous with respect to the locus of learning. Kourtzi et al. (2005) took the lack of a change in BOLD signal in early visual areas to indicate the absence of learning in early visual areas. However, this null effect might be expected if a large number of neurons responded weakly to a given stimulus prior to learning, and a small set of neurons responded highly selectively and more strongly after learning; that is, the null effects are consistent with the hypothesis that the brain developed sparse representations in these regions. Indeed, a trade-off between number of active cells and response intensity has been used to explain why the BOLD signal is often insensitive to on-line processing differences to familiar and unfamiliar (novel) words in the visual word form area (Glezer et al., 2009).

Accordingly, the null effect of BOLD signal in V1, V2, VP or V4 does not rule out learning in these areas, and the reduced BOLD signal in LOC may reflect knock-on effects from learning in one or more of these early visual areas [scenario (e) in Fig. 1]. Alternatively, the reduction of activation in LOC may reflect learning in higher visual areas, with feedback reducing the activation of LOC [as in scenario (f)]. Or, some combination of the above might explain the observed pattern of BOLD responses.

The same ambiguity applies to studies reporting selective changes in activations in low-level visual areas. Yotsumoto et al. (2008) examined gradual changes in V1 activation as a result of 14 training sessions in a texture discrimination task. Training was restricted to only one quadrant of the visual field, and BOLD responses were recorded before the start of training, and following one, six and 14 sessions. Regions of interest included V1, V2, and V3/VP. Behaviourally, performance improved from sessions 1–6, and then remained constant. The imaging data, by contrast, showed increases in BOLD signal in V1 at times 2 and 3 compared to time 1, and a decrease at time 4, without any significant differences in the other regions of interest. Furthermore, both the behavioural and BOLD changes were largely location specific. That is, when participants were tested outside the trained quadrant, behavioural learning and BOLD signals were little affected by training. A two-stage model of perceptual learning was proposed to account for this pattern of findings: an initial stage marked by an increase in the number and strength of synaptic connections in V1 (accounting for the increased BOLD signal), and a second stage, starting after behavioural performance levels have saturated, marked by downscaling of synaptic connections in V1, leaving only the most critical synapses for task execution (accounting for the decreased BOLD signal).

Although this explanation is plausible, the conclusions are not warranted by the data. Indeed, as above, the lack of a BOLD change following training in areas V2 and V3/VP does not rule out the occurrence of learning in these areas. Furthermore, the changes in V1 BOLD signal could reflect learning in other brain areas that do not show a change in BOLD activation (or were not included in the analysis), with knock-on effects due to feedback to V1 [scenario (f) in Fig. 1]. In fact, top–down on-line processing effects can impact on BOLD signal in many ways. For example, functional connectivity analyses carried out on perceptual learning in a texture discrimination task showed that performance in an untrained condition was associated with functional coupling between early visual cortex and areas in frontal and parietal cortex (Schwartz et al., 2002). This was taken as evidence of top–down feedback mechanisms of spatial attention in the initial stages of learning, and may also play a role in V1 BOLD changes with training.

What is to be made of the fact that Yotsumoto et al. (2008) found that changes in both behaviour and BOLD signal were selective to training location? Does this provide evidence regarding the locus of learning? The behavioural results support the claim that learning has occurred within retinotopic cortex where receptive fields of the neurons are restricted to a given quadrant of the visual field, but such cortical areas extend well beyond V1, including V2, V3/VP, and perhaps V4. The critical point to make is that the BOLD signal is not providing any additional information regarding the locus of learning. That is, the imaging results do not restrict the learning to V1, and indeed, the learning may have occurred outside V1, with feedback from higher levels of retinotopically organized visual cortex (e.g., V2, V3/VP, and V4).

We do not mean to imply that the data reported by Kourtzi et al. (2005) and Yotsumoto et al. (2008) are uninformative. Indeed, the studies may provide important constraints for theories of visual learning. For example, Kourtzi et al. ’s finding that the BOLD signal increased in the LOC following training in the low-salience condition and decreased following training in the high-salience condition demonstrates that the learning in these different conditions has impacted on the on-line processing in different ways. These findings need to be explained. Our point is simply that these findings do not provide powerful constraints regarding the locus of learning. Other tools are needed.

In sum, the standard approach to study the locus of learning with fMRI has been to look for changes in BOLD over the course of training. Any changes (increases or decreases) are thought to reflect changes in on-line processing in the corresponding brain area, which in turn is taken to reflect the locus of learning. We accept the first link in this argument, but claim that the second one is flawed. That is, changes in on-line processing can reflect learning local to a brain area, or alternatively, reflect knock-on effects of learning that has occurred elsewhere.

4. Connectivity analyses using BOLD signal to study the locus of learning

As we have shown, there are two main limitations with using fMRI to study loci of neural plasticity: (1) the indirect link between the BOLD signal and specific molecular, cellular or systemic changes underlying learning; and (2) the potential for strong dissociations between the BOLD signal and the locus of neural plasticity due to the bottom–up and top–down feedback connections involved in on-line processing (as depicted in Fig. 1). We now investigate whether other fMRI-based methods can address these limitations more satisfactorily.

Connectivity analyses adopt a reasonable assumption that neural substrates employed in a cognitive task are usually spread across brain regions. One version of this technique, namely, functional connectivity analysis, investigates
changes in the correlations between BOLD activations across brain regions before, after, and sometimes during learning, and often in relation to behavioural performance. For example, Mukai et al. (2007) examined functional connectivity changes during one session of training in a contrast discrimination task. Higher correlations between visual areas (including V1, V3 and V4) and attention-related areas (mainly intraparietal sulcus) were found in good learners in comparison to poor learners. According to Mukai et al. their findings suggest “enhanced integration of information processing between attention-related regions and early visual areas” in good learners (p. 11410). These results are interesting, but do not speak to the issue of the locus of neural plasticity that supports the enhanced integration. It could reflect learning in retinotopic cortex resulting in more efficient low-level processing that would support interactions with the attentional control areas, or the locus of learning could be in intraparietal sulcus with top–down feedback to visual cortex, or perhaps some other area that interacts with the areas showing changes in BOLD activations.

Another version of this technique, so-called effective connectivity analysis, uses models of causal relationships between brain areas and pathways connecting regions to guide correlations in BOLD activation changes (Horwitz et al., 2005; McIntosh, 1999). For example, Büchel et al. (1999) found decreases in regional BOLD signal activations after associative learning and first attributed this finding to refined selectivity of neurons within these areas. However, effective connectivity analysis suggested that the decreased level of activation is better explained by the increased effective connectivity between active regions.

But despite the potential advantage of connectivity analyses, the fundamental problems with using fMRI data to make inferences regarding the locus of learning remain. The analyses provide evidence that the two areas are more strongly connected after learning, but whether this reflects learning in one region or another remains unclear given the link between BOLD and locus of plasticity is so weak. Indeed, all the learning scenarios depicted in Fig. 1 might be expected to impact on effective connectivity between levels, but the locus of learning varies.

In short, both functional and effective connectivity analyses provide new ways to explore metabolic changes associated with learning. Nevertheless, their reliance on fMRI data still makes them susceptible to the two major problems outlined above.

5. Imaging without BOLD signal in the study of the locus of neural plasticity

Now we briefly turn to imaging techniques which do not rely on BOLD signal and examine whether they can provide more direct measures of the loci of neural plasticity.

Voxel-based morphometry (VBM) provides a measure of structural changes in the brain and has often been used in research on learning by imaging modifications of grey matter density. For example, extensive navigation experience has been associated with a significant increase in grey matter volume in posterior hippocampi of taxi drivers compared to controls (Maguire et al., 2000; 2006). The volume increase correlated positively with time spent in the profession. Other studies found changes in grey matter volume as a result of learning to juggle (Draganski et al., 2004) and after 2 weeks of short daily practice in mirror-reading (Ilg et al., 2008).

One limitation of this approach is that the exact nature of neural changes detected by VBM is not well understood (Draganski and May, 2008). Modifications of grey matter volume will sometimes be associated with neural plasticity (e.g., changes in numbers of neurons and structure of dendrites), but other changes in grey matter may reflect modifications less relevant to learning (e.g., changes in fine vasculature). Perhaps the main limitation of this technique is that it can only pick up learning that results in gross structural modifications (Thomas et al., 2010). This may exclude most learning studied in the lab.

Diffusion tensor imaging is another method that measures structural changes in the brain and can trace white matter tracts (Bandettini, 2009). Recently this technique has been used to investigate learning. For example, Scholz et al. (2009) reported changes in white matter of intraparietal sulcus after six weeks of juggling, and this was attributed to changes in axon myelination or axon density. But again, there are a number of limitations of this method at present. First, it is not clear how axon myelination relates to learning. Second, much like connectivity analyses with fMRI, this technique provides evidence that two brain areas are more strongly connected following learning, but the locus of learning is unclear (unless increased myelination is itself the locus of learning). Third, as with the case of VBM, this technique is only able to identify learning following substantial practice that alters the gross anatomical structure of the brain. Whether this technique could pick up learning following 20 min practice in a perceptual task (as common in many perceptual learning tasks) is unclear.

Finally, magnetic resonance spectroscopy (MRS) detects concentrations of metabolites and neurotransmitters in tissue. With regard to neural plasticity, concentrations of GABA (γ-aminobutyric acid) are of particular interest because of their mediating role in synaptic transmission underlying long-term potentiation (LTP) and long-term depression (LTD) as central mechanisms of learning (Collingridge et al., 2004). Applied to imaging research on learning, a significant decrease in GABA concentration in primary sensorimotor cortex has been found after 30 min of force tracking training involving contralateral hand movement (Floyer-Lea et al., 2006). The decrease in GABA was associated with improvement in performance and most likely reflects facilitation of LTP through reduction in GABAergic inhibition. The opposite effect has been documented in another study using MRS (here in conjunction with transcranial magnetic stimulation), in which increased GABA concentrations have been linked to inhibition in synaptic transmission associated with long-term depression in primary motor cortex (Stagg et al., 2009). The link between the locus of learning and the MRS imaging data in these studies seems more direct than in any of the functional and structural MRI methods reviewed. Given the lack of studies on learning using MRS, it is difficult to judge to what extent this method is subject to the problems identified above. One of the most notable technical shortcomings of the method is its poor spatial resolution. Still, among the methods
considered here, it seems (to us) the most promising approach to studying the locus of neural plasticity to date.

6. Future directions

We have argued that fMRI studies of learning are subject to methodological limitations that do not apply to fMRI studies of on-line processing. In our view, these limitations significantly undermine conclusions derived from existing fMRI studies concerned with the locus of learning.

We suggest the following recommendations for future imaging studies designed to identify the brain regions involved in learning perceptual and other skills. First, fMRI studies need to distinguish outcomes of learning reflected in on-line processing from the locus of learning (neural plasticity), because BOLD is more sensitive to the former than the latter. As shown in Fig. 1, the locus of these two processes can be sharply dissociated. Second, if fMRI research on learning is to progress, more focus on studies investigating links between BOLD signal changes and underlying neural mechanisms of learning are needed. If strong links can be found, then other techniques should be favoured. Third, two basic questions should guide the selection of an imaging method when studying learning in a given task, namely, What are the most likely neural changes associated with learning in a particular task? and what is the most direct method to detect these changes? These questions are the key, as little progress is possible in identifying the brain regions that support perceptual learning (and learning more generally) without direct and sensitive measures of the underlying processes that mediate learning.

Acknowledgments

Authors would like to thank Guillaume Thierry and Paul Mullins for useful comments on an early draft of this article.

References


Coltheart M. What has functional neuroimaging told us about the mind (so far)? Cortex, 42(3): 323–331, 2006.


