Chapter 4

When does Visual Attention Modulate fMRI Activity in High-level Visual Areas?

4.1 Introduction

The work presented thus far demonstrates that the visual system is well equipped to handle the stimuli it encounters on a daily basis. As we have argued, rather than relying on laborious attention-dependent processing, the underlying architecture of the visual system capably supports straightforward computations of natural objects and scenes. It has been suggested that a possible basis for these computations could be the neuronal selectivity observed in higher-level brain areas for natural stimuli (VanRullen, Reddy et al., 2004). Neurons in the anterior part of the infero-temporal cortex in monkeys and the medial temporal lobe in humans have been observed to be highly selective to natural categories such as animals, faces, foodstuffs, buildings, and so on ((Gross, Rocha-Miranda et al., 1972; Perrett, Rolls et al., 1982; Desimone, Albright et al., 1984; Kreiman, Koch et al., 2000a; Quian Quiroga, Reddy et al., 2005); see also Chapter 5 of this thesis), and these built-in preferences for different types of natural stimuli could ensure that the corresponding stimuli reach perception in the near-absence of attention, resulting in high levels of performance on the discrimination tasks.

However, several fMRI studies have shown that BOLD activity for unattended objects is decreased or even abolished in high-level areas specifically encoding these stimuli (Wojciulik, Kanwisher et al., 1998; O'Craven, Downing et al., 1999; Vuilleumier, Armony, Driver, & Dolan, 2001; Pessoa, McKenna et al., 2002; Marois, Yi, & Chun, 2004; Yi, Woodman, Widders, Marois, & Chun, 2004). In these studies, the BOLD signal measured in the fusiform face area (FFA) or the parahippocampal place area (PPA), (areas that typically show enhanced activation to images of faces and places, respectively) was shown to be significantly reduced in the absence of attention.[1] These results appear to contradict the observation that subjects are efficiently able to categorize natural stimuli when attention is not fully available. For if the cortical areas that supposedly support the processing of these stimuli show significantly lower levels of activity in the near-absence of attention, then how is it that the corresponding behavioral performance remains unaffected?

In this chapter, in an attempt to reconcile these two lines of evidence, we studied the effects of attentional modulation on face processing in the FFA, a region in the fusiform gyrus that responds strongly to images of faces compared to other non-face stimuli (Puce, Allison, Gore, & McCarthy, 1995; Kanwisher, McDermott et al., 1997; Kanwisher, Stanley et al., 1999). We employed the dual-task paradigm in a 3.0T fMRI scanner while subjects performed a face-gender discrimination task. We must note here that an animal or vehicle discrimination task could have served just as well in principle. However, in practice, faces turn out to be a very convenient set of stimuli to use in fMRI studies since the FFA shows very specific responses to faces, and can be readily localized (Kanwisher, McDermott et al., 1997; Kanwisher, 2000).

Previously, fMRI studies have investigated the effects of attentional modulation on face processing as paradigmatic for object processing in general (Wojciulik, Kanwisher et al., 1998; O'Craven, Downing et al., 1999; Vuilleumier, Armony et al., 2001; Pessoa, McKenna et al., 2002). Most of these studies compared activity in the FFA in two conditions: when the faces were fully attended and task relevant, and when

---

[1] Attention was manipulated using various methods: e.g. the attentional blink paradigm, face/house matching tasks, or reporting on the orientation of peripherally presented bars, which were displayed along with the faces.

the faces were fully ignored (and task irrelevant). A significant drop in BOLD activity has generally been observed between these two conditions, and the difference was attributed to attentional modulation. However, because subjects were instructed to make a behavioral report on the faces in the first, but not in the second condition, it is unclear whether the observed drop in activity is solely due to attentional modulation, or if it can also be explained by a change in the behavioral relevance of the faces.

An advantage of the dual-task paradigm is that, in addition to the two conditions used in previous studies (faces attended vs. faces ignored), it includes a third condition in which attention is spatially focused away from the faces (on an attentionally demanding letter discrimination task), yet they remain task relevant. Thus we are in a position to separately examine the effects of focal attention and task relevance on activity in the FFA. In agreement with previous reports, we observed a significant drop in activity when the faces were unattended and behaviorally irrelevant. However, when the spatial focus of attention was not on the faces but they remained behaviorally relevant, activity was left unaffected. This high level of activity in the near-absence of focal attention might constitute the neural basis of the surprising behavioral performance obtained in the dual task condition.

4.2 Methods

4.2.1 Behavioral training

Twelve paid volunteers, aged 20–30 participated in dual-task face-gender discrimination experiments. All experiments were conducted according to the guidelines of the Institute's committee for protection of human subjects. The details of the paradigm are described in Chapter 2. For the current set of experiments, six "trained" subjects were trained on the two single-task conditions (letter and face-gender discrimination) as well as the dual-task condition, while six "untrained" subjects were trained on only the two single-task conditions. The entire training procedure typically lasted between 5 and 10 hours on consecutive days for each subject and was carried out outside the scanner on a Silicon Graphics computer. The details of the training procedure were identical to those described in Chapter 2. Over the group of 12 subjects, the SOAs varied between 133–160 ms and 173–240 ms on the face and letter tasks, respectively.

The "trained" subjects also performed a control experiment outside the scanner. In this experiment, the peripheral task involved discriminating a vertically bisected disk with red and green halves (equated for gray values) from their mirror image. Each disk was masked by a disk divided into four red and green alternating quadrants. Subjects received equal amounts of training on this task (central, peripheral, and dual blocks) as on the face-gender discrimination task; the tasks were matched for difficulty such that average single-task performance was around 80%.

The behavioral performance of subjects on the dual-task experiment is reported below both in terms of percent correct, as well as the discriminability index, d' (a measure based on signal detection theory, which gives a response bias-free measure of performance).

Normalized performance values reported in Figure 4.2 are calculated by a simple linear scaling of the mean value of each participant's performance. The scaling mapped the mean single-task performance to 100%, leaving chance at 50%:

Normalized performance = $0.5 + 0.5.[(P_2-0.5)/(P_1-0.5)]$, where P_2 and P_1 refer to performance in the dual-task and single-task conditions respectively.

d’ was calculated for the letter and gender discrimination tasks by estimating the proportion of hits and false alarms on each task. Note that d’ is a measure of the distance between the noise and signal+noise distributions and assumes that the noise follows a Gaussian distribution with a fixed variance.

4.2.2 fMRI sessions

The twelve subjects performed the face-gender, dual-task experiment in a 3.0 T whole-body Siemens scanner at the California Institute of Technology. Stimuli were presented on a Macintosh laptop and viewed through optical goggles (Resonance Technologies, Northridge, CA, VisuaStim XGA. 800 x 600 resolution at 60 Hz). The size and eccentricity of the stimuli in the goggles was the same as that during training. T2* weighted Echo-planar images (TR=2s, TE=30ms, FA=90°, FOV=210mm, 31X3 mm interleaved axial slices) were acquired using the whole-head coil and an in-line motion correction sequence (Thesen, Heid, Mueller, & Schad, 2000). The slices were positioned to cover the temporal lobe. Each functional run consisted of 186 volumes; the first two volumes were discarded. A 12-minute high resolution (1x1x1mm) whole-head T1-weighted MPRAGE sequence (T1=1.5s, TI = 0.8s, TE=3.05ms, 176 sagital slices) was also acquired for each subject.

Localizer task

FFA and PPA

Subjects were first presented with sequences of images in order to localize the FFA and PPA (Kanwisher, McDermott et al., 1997). The design consisted of alternating 30 s blocks of color faces and outdoor scenes (approximately 9 x 13.5° of visual angle), obtained from the Vision Lab at Caltech (http://www.vision.caltech.edu/html-files/archive.html) presented for 1 s each at fixation. The faces were different from the ones used in the dual-task experiment. Each run lasted for 4 minutes, and subjects participated in 3 runs. The FFA and PPA ROIs were localized in each individual by contrasting average brain activity in face versus scene blocks. Due to different signal-to-noise ratios, there was some variation in the threshold used for defining the ROIs across subjects. The average t value that defined the ROI was $t(6) = 4.1 \pm 0.4$; $p < .005$. The mean Talairach coordinates of the FFA and PPA (right FFA: $x = 38 \pm 1$mm, $y = -50 \pm 2$ mm, $z = -13 \pm 1$mm; left FFA: $x = -40 \pm 1$mm, $y = -51 \pm 2$mm, $z = -15 \pm 2$mm; right PPA: $x = 24 \pm 2$mm, $y = -42 \pm 2$mm, $z = -6 \pm 1$mm; left PPA: $x = -28 \pm 1$mm, $y = -46 \pm 4$mm, $z = -9 \pm 2$mm) are consistent with previous reports (Kanwisher, McDermott et al., 1997; Epstein, Graham, & Downing, 2003). The mean number of voxels over all subjects was 28 ± 3 and 67 ± 29 in the FFA and PPA, respectively. The FFA and PPA localized for one subject are shown in Figure 4.1. In a separate localizer session, 4 of the 6 subjects were presented with a rapid sequence of faces and scenes in the periphery at 5Hz. The faces and scenes were presented at the same locations as during the dual-task experiment (see below). At each peripheral location, the sequences of faces and scenes were presented for a total of 2.5 s, with each face or scene being shown for 200 ms.

LOC

The Lateral Occipital Complex (LOC) in both hemispheres was localized for the 6 untrained subjects by contrasting images of grayscale objects obtained from (Murray & Wojciulik, 2004) with scrambled versions of these images. The design and analysis of these localizer runs were otherwise identical to those for identifying the FFA and PPA as described above.

Dual-task paradigm

The fMRI experiment was based on a block-design. Subjects performed 10 runs of the dual-task paradigm in the scanner using the parameters described in Chapter 2 except for the following modifications. In a single run, 2 blocks each of the central, peripheral- and dual-task conditions were presented in a randomized counterbalanced order, with each block consisting of 16 trials. On average each block lasted approximately 40 s and was followed by a 20 s blank interval. Subjects made their behavioral report using a 4 button response box. As in Chapter 2, the peripherally presented face subtending approximately 2.5 degrees of visual angle appeared at a random location on one of the four edges of an imaginary rectangle. This imaginary rectangle measured 8 degrees by 10 degrees of visual angle. The eccentricity of the faces was thus 4 to 5 degrees, (minimum = 4, maximum = 6.4 degrees (on the diagonal)).

Data analysis

For each subject, images were co-registered between runs using a Linear Image Registration Tool (FLIRT) (Jenkinson, Bannister, Brady, & Smith, 2002). Analyses were conducted using custom Matlab software (Mathworks). For each subject, the time-course in the 3-task conditions was calculated by averaging the BOLD signal change in all voxels of the relevant ROI over all repetitions of each condition (after removing linear drifts over the course of each run). These time-courses were collapsed across hemispheres and subjects. This is the data shown in Figures 4.4a, c and 4.6a, c and 4.8a, c. The time-courses for each subject were then integrated over the duration of the 40s block (shifted by 7s to account for the delay in the hemodynamic response) and normalized by dividing by the mean activation over all 3 conditions (group data shown in Figures 4.4b, 4.6b, and 4.8b. Statistical analyses (ANOVA and paired *t*-tests corrected for multiple comparisons using Scheffe’s method) were performed on these values.

4.3 Results

4.3.1 Behavioral performance

We used the dual-task paradigm described in Chapter 2 to examine the effects of attentional manipulation on brain activity in the FFA and the PPA. A typical trial from the experiment is shown in Figure 4.2a. In all trials, both the face and letters were displayed as shown, and depending on instructions, subjects could perform one of three tasks : (1) In the peripheral face-gender task condition, subjects had to report whether the face was male or female. Since attention was focused on the faces, and subjects made a behavioral report on them, this condition corresponds to the “attended” condition of most previous studies. (2) In the central letter task condition, subjects reported whether the letters were all the same, or if one differed from the other four. In this condition, attention is focused on the letters and away from the faces, which are task irrelevant. Thus this task is similar to the “unattended” condition in previous studies. (3) In the critical dual-task condition, subjects had to perform on both tasks simultaneously. As we have shown previously ((Reddy, Wilken et al., 2004); Chapter 2) and discuss below, in this condition, focal attention must be focused away from the faces although subjects are required to make a behavioral report on them. Thus, this condition allows us to dissociate the effects of attentional manipulation and task relevance (Fig 4.2b).

Six subjects who had been previously trained on the face-gender, dual-task paradigm, performed 20 blocks each of the central, peripheral, and dual-task conditions in the scanner. Each 16-trial block lasted 40 seconds and was followed by a blank display for 20 seconds. The behavioral results for these subjects on this task are shown in Figure 4.3a. As reported previously (Chapter 2 and (Reddy, Wilken et al., 2004)), participants performance on face-gender discrimination in the dual-task condition was comparable to that obtained in the single-task condition (mean performance ± s.e.m. normalized to single-task performance : 92.5 ± 2.0%; difference in performance between single- and dual-task conditions was significant for 3 of the subjects ($p<.05$)). The performance of these subjects was comparable to the performance of the six subjects reported in the first experiment of Chapter 2 (performance on central letter task, $p=0.79$; performance on face-gender task, $p=0.69$). Thus, a high level of performance is achieved even when the spatial focus of attention is shifted away from the faces. This conclusion also holds if we use d’ as a response bias-free measure of performance: d’ in the single- and dual-task conditions were on average 1.8 ± 0.5 and 1.4 ± 0.3 (mean ± s.d), respectively, and were not significantly different ($p=.15$). Thus a change in strategy between these two conditions is unlikely to explain the good performance we observe in the dual-task condition. Similarly d’ values were not significantly different between single and dual-task conditions of the central task ($p=0.1$). Eye movements do not play a major role in achieving this performance since the peripheral faces are presented only briefly and at random locations. Furthermore, control experiments with an eye-tracker have allowed us to verify the absence of eye movement contribution to this performance (Chapter 2).

An obvious concern that arises, however, is whether the central letter discrimination task is sufficiently demanding to engage the focus of attention away from the periphery. This concern can be addressed by verifying that performance on known, attentionally demanding tasks suffers under the same dual-task conditions. As shown in Figure 4.3b, performance for five of the subjects (the remaining subject was not available for testing) on a bisected disk discrimination task in the periphery falls dramatically when performed concurrently with the central letter task, even though their performance on this task was comparable to face-gender performance when both tasks were performed alone. The difference in d’ values on this task between peripheral (1.9 ± 0.2) and dual-task (0.25 ± 0.1) conditions was significant ($p<0.00005$) thus demonstrating that the central letter task did effectively engage the focus of attention away from the periphery. As before, we thus define focal attention operationally as that resource that, when engaged by the central T/L discrimination task, is unavailable to the periphery, consequently impairing performance on some concurrent tasks (such as bisected disk discrimination but not face-gender discrimination or natural scene categorization. (See also (Braun & Julesz, 1998; Lee, Koch et al., 1999; Li, VanRullen et al., 2002)).

4.3.2 fMRI activity: main result

The FFA and PPA (Epstein & Kanwisher, 1998) of each subject were isolated in separate localizer scans run before the dual-task sessions by contrasting brain activity in blocked presentations of faces and scenes presented centrally. The BOLD activity during the dual-task experiment was analyzed separately in the isolated FFAs from the left and right hemispheres. Since similar results were observed in both hemispheres, in the data presented here, the FFAs were collapsed across hemispheres. Figure 4.4a shows the raw time course of BOLD activity in the FFA during the three experimental conditions. Consistent with previous reports, we observed a significant decrease in activity when the spatial focus of attention was removed from the faces and they were ignored (peripheral task condition vs. central task condition, [$F(2,15) = 11.89$, $p = 0.0008$], Figure 4.4b). Remarkably, in the dual-task condition, when attention was not focused on the face, but a gender-specific response was still required, activity was as high as in the peripheral-task condition. Thus, we observed a dissociation between the effects of attentional manipulation *per se* and task relevance.

To rule out the possibility that BOLD activity was saturated in the peripheral and dual-task conditions (thereby accounting for the similar levels of activity in these conditions), we compared peripheral-task activity with that evoked by the face stimuli during a localizer experiment in the FFA when faces were presented peripherally at 2Hz, in both hemispheres of our subjects. We observed significantly lower levels of activity in the former condition compared to the latter (mean ± s.e.m.: 1.0 ± 0.1 vs. 1.9 ± 0.2 % signal change, t(22) = 2.9, $p < 0.01$). Note that this result is not unexpected: it is known that peripherally presented faces (peripheral-task condition) activate the face regions significantly less than centrally presented ones (localizer runs) (Levy, Hasson, Avidan, Hendler, & Malach, 2001). The frequency of presentation of faces (once per trial) was also considerably less than in the localizer scans. On the other hand we cannot rule out the possibility that a subset of neurons within these voxels, coding specifically for our peripheral face stimuli, could have been saturated in the dual-task experiment. To control for this possibility, we measured the activation in the FFA to peripherally presented faces and scenes (see Methods for details of the second localizer experiment). The faces were presented at 5Hz, and we observed significantly higher activity to these rapidly presented peripheral faces (2.1 ± 0.4 % signal change; $p < .05$) compared to activity in the peripheral-task condition. Thus the similar levels of activity observed between the dual and peripheral-task conditions truly mirror the experimental manipulation rather than saturation in the BOLD signal.

In contrast to a recent study that reported total absence of activation to unattended and ignored faces (Pessoa, McKenna et al., 2002), we observed decreased but significant levels of BOLD activity in the FFA in the central-task condition, presumably in response to the peripheral and task-irrelevant faces (Figure 4.4 a,b). We estimated the amount of activation elicited by the faces in this condition by examining the effect of removing the faces altogether. In a control experiment, three of our subjects performed the letter discrimination task exactly as in the central-task condition with the exception that the faces were not presented in the periphery. We observed a large (67%) and significant ($p < .05$) drop in FFA BOLD activity when the faces were not presented compared to the central-task condition when the faces were present although unattended and fully ignored (Figure 4.5b). Thus, we find that even unattended and ignored faces produce significant ($p < .05$) levels of FFA activation, in agreement with previous reports (Wojciulik, Kanwisher et al., 1998; Vuilleumier, Armony et al., 2001). In contrast to this relatively large effect of eliminating the influence of face stimuli, removing the letter stimuli from the peripheral task condition resulted in a much smaller (27%) drop in FFA activity (Figure 4.5c).

Undoubtedly, the dual task condition is more demanding than either of the two single tasks, since subjects have to perform two tasks simultaneously. Therefore, subjects might compensate for this perceived increase in difficulty by a more aroused state resulting in a widespread increase in activity in the dual-task condition. To determine whether this effect accounts for the high level of activity obtained in the dual-task condition, we looked at activity in the PPA region that had been identified in the localizer runs. This region is more strongly activated by spatial layouts than by faces (Epstein & Kanwisher, 1998). As shown in Figure 4.4c, no differential effect was present in the PPA ($p > .05$), thus ruling out the possibility that non-specific arousal effects are responsible for the observed results.

4.3.3 fMRI activity: the role of training

Even for simple tasks such as color or orientation discrimination, achieving good performance in the dual-task paradigm is contingent on substantial training (Braun & Julesz, 1998; Li, VanRullen et al., 2002). We thus wondered whether the high level of activity in the dual-task condition was merely a consequence of training. In other words, can everyone profit from this ability to process faces outside the focus of attention, or is it only acquired after extensive training? Furthermore, since untrained subjects usually show a drop in behavioral performance in the dual-task condition, we were also interested in determining if FFA activity would mirror this decrease. Therefore, in a second experiment, six subjects, untrained in the dual-task condition performed the face-gender dual-task experiment in the scanner.

Six subjects who had never performed in any dual-task condition before (but had been exposed to both single tasks) were tested in the scanner under exactly the same conditions as in the previous experiment. As expected, their behavioral performance on face-gender discrimination dropped considerably when the focus of attention was shifted away from the faces: for this group of subjects, average performance in the dual-task condition was 73.7 ± 3.0% of performance in the single-task condition (difference in performance between single- and dual-task conditions on the peripheral task was significant for all six subjects, $p<0.009$; Figure 4.3c). On average, the corresponding d' dropped significantly ($p=.008$) from 1.3 ± 0.4 in the single-task to 0.6 ± 0.3 in the dual-task conditions.

Figure 4.6 shows the activity in the FFA for these six subjects. As in the data for Figure 4.4, levels of BOLD activity in the dual-task and peripheral-task conditions were similar, and a significant reduction was only observed in the central-task condition [$F(2,15) = 9.46$, $p = 0.002$]. Thus, despite the substantial drop in behavioral performance in the dual-task condition, the corresponding BOLD activity did not decrease compared to the single-task level. This implies that training is not a necessary condition for maintaining high levels of FFA activity in the absence of focal attention.

4.4 Discussion

Our results demonstrate a dissociation between the effects of focal, spatial attention and behavioral relevance. Manipulating the spatial focus of attention alone did not significantly affect the BOLD signal in the FFA as long as the face stimuli remained relevant to the task at hand. At first glance, this result may appear to be at odds with a number of other studies that have reported a significant decrease in FFA BOLD activity as a result of attentional manipulation of faces (Wojciulik, Kanwisher et al., 1998; O'Craven, Downing et al., 1999; Vuilleumier, Armony et al., 2001; Pessoa, McKenna et al., 2002). In fact, the two experimental conditions measured in most of these studies roughly correspond to our peripheral and central task conditions: subjects either perform a task involving the faces or ignore them completely. And, in agreement with previous findings, we observed significantly distinct levels of BOLD activity between these two conditions. However, by introducing a third condition in our paradigm (the dual-task condition), we were able to separate the effects of two factors (spatial attention and task relevance), which were confounded in previous reports. Our data shows that these factors indeed have distinct effects on the BOLD signal in the FFA. Thus, rather than being at variance with previous reports, these results, together with earlier findings, reveal a more complete picture of how focal attention and behavioral relevance interact in higher levels of visual cortex.

4.4.1 Focal attention and behavioral Relevance

Our interpretation here relies on a clear distinction between focal, spatial attention and top-down behavioral relevance. But how exactly do we define behavioral relevance? It is obvious from our data that there are two factors at play in our paradigm. We have operationally defined focal attention as the resource that is needed to solve the central task. Similarly we use the term “behavioral relevance” as an operational definition for the distinction that arises in our paradigm as a result of giving subjects different face-based task instructions in the three conditions. It could be argued, that in fact, what we term top-down behavioral relevance is a form of top-down and/or feature-based attention. Indeed, previous reports have shown that paying attention to a particular feature increases feature-related activity throughout the visual field (Treue & Martinez Trujillo, 1999; Saenz, Buracas et al., 2002). We believe that our interpretation is not incompatible with such alternative definitions, provided one assumes that face-gender constitutes a “feature” for the visual system.

Although we observed that the FFA BOLD activity was primarily dependent on behavioral relevance, irrespective of the current focus of attention, we do not wish to imply that behavioral relevance alone is the dominant factor affecting neuronal responses. Our paradigm does not allow us to examine the reverse interaction, namely how focal attention would affect processing when the stimuli are not task relevant. It is entirely possible that focusing attention on task irrelevant stimuli enhances their representation in cortex in a manner similar to when poorly attended stimuli are made behaviorally relevant (central vs. dual-task conditions). Although this issue has not been addressed explicitly, indirect evidence comes from studies in which manipulating the attentional load results in a significant modulation of BOLD activity in response to task-irrelevant stimuli in several brain areas (Rees, Frith, & Lavie, 1997; O'Connor, Fukui, Pinsk, & Kastner, 2002; Pinsk, Doniger, & Kastner, 2004; Yi, Woodman et al., 2004). Thus, focal attention and behavioral relevance might very well have interchangeable effects on brain activity, and it remains to be seen how these effects add up and/or interact in the visual cortex (see for example, (Treue & Martinez Trujillo, 1999)).

4.4.2 Effects of training in the FFA

What is the specific role of training in our paradigm? Our data shows that, within each group of subjects, FFA activity obtained in the dual and peripheral-task conditions was comparable (Figures 4.4 and 4.6 a, b). Additionally, between the two groups of subjects, the BOLD activity in the central, peripheral, and dual-task conditions did not differ significantly (p=0.3, p=0.3, and p=0.7, respectively). In terms of behavioral performance, for both groups of subjects, the information provided by this level of activity was sufficient to yield good performance, at least in the peripheral-task condition (presumably FFA activity can constitute a basis for behavioral decisions regarding face stimuli). However, in the dual-task condition, the same level of FFA activity only seemed to benefit the trained subjects. Thus, it appears that although information was similarly available to both sets of subjects, only the trained subjects were able to use it effectively. Although we cannot discount the (unlikely) possibility that the FFA is not necessary for face-gender discrimination tasks, it is possible that the BOLD signal is too crude a measure to show a difference in activation between trained and untrained subjects. Alternatively, rather than increasing BOLD activity in the FFA, training could facilitate a more efficient use of the relevant activity, possibly by minimizing competitive bottlenecks that arise (at stages later than the FFA) from processing two tasks simultaneously (Allport, 1980; Duncan, 1980b; Pashler, 1984, 1994).

4.4.3 The fMRI signal and behavioral performance in the FFA

It has been previously reported that the fMRI signal correlates with behavioral performance during object recognition tasks (Grill-Spector, Kushnir, Hendler, & Malach, 2000; Bar, Tootell et al., 2001). In the study by Grill Spector et al., various objects were presented at very short SOAs (20-500ms) and subjects’ performance at correctly recognizing these objects was correlated with changes in the BOLD signal. Bar and colleagues presented masked objects for 67 ms and correlated the BOLD signal with subjects’ confidence in reporting the identity of the objects. Both studies report that higher levels of behavioral performance were correlated with larger increases in the BOLD signal.

However our data reveal that for our subjects there is no apparent correlation between behavioral performance and BOLD activity in the FFA. Whereas we observed significantly lower behavioral performance for our untrained subjects compared to the trained subjects, both groups had comparable levels of FFA BOLD activity in the dual-task condition, as discussed above. Over the groups of trained and untrained subjects, the correlation coefficients (r^2) between the average FFA BOLD activity of a subject and their average face-task performance (over all face-task blocks) were 0.04 and 0.08, respectively. At the level of individual subjects, additional analysis revealed that on each face-task block, there was no correlation between BOLD activity in the FFA and the corresponding behavioral performance ($r^2 = 0.001$ on average, and $r^2 < 0.13$ in all cases; Figure 4.7 shows this data for the ‘trained’ subjects). Thus, while FFA activity might be necessary for good behavioral performance in our gender discrimination task, it is not sufficient. One possible difference from earlier studies (Grill-Spector, Kushnir et al., 2000; Bar, Tootell et al., 2001) lies in the procedure used to modulate performance. These studies used backward masking with short SOAs (20–500 ms, compared with ~160 ms for our paradigm) to limit stimulus visibility. It is possible that the attentional manipulation in our paradigm limited discriminability (and thus behavioral performance) without affecting the visibility of the faces (although we did not directly test for stimulus visibility). When male faces are confused with female faces at the same level of visibility (and vice-versa), no net difference in FFA activation is expected.

4.4.4 Visual responses in the FFA

As we mentioned previously, faces are a convenient set of stimuli to use in fMRI studies because the regions (Kanwisher, McDermott et al., 1997; Halgren, Dale et al., 1999; Haxby, Hoffman et al., 2000; Hoffman & Haxby, 2000) responding to face stimuli can be easily identified. However, although it has consistently been shown that these regions respond more to faces than to other objects, whether or not these areas are specialized for face processing is a heavily debated matter. In particular, fMRI studies have observed significant responses to other types of natural stimuli, including animals in the FFA ((Chao, Martin, & Haxby, 1999; Ishai, Ungerleider et al., 1999), but see also (Kanwisher, Stanley et al., 1999)). Besides the stimulus category, other factors such as the subjects' expertise with the relevant stimuli are thought to play a role in determining the visual properties of these areas as well (Gauthier, Tarr et al., 1999; Gauthier & Logothetis, 2000; Gauthier, Skudlarski et al., 2000; Gauthier & Nelson, 2001). In this context, our results suggest that the behavioral relevance of the stimuli must also be taken into account when considering the visual response of the FFA.

Another factor that is thought to influence processing in the FFA (and other higher level cortical areas) is related to the eccentricity at which the stimuli are presented. Specifically, it has been suggested that the FFA is associated with tasks that require a great degree of visual acuity (such as processing facial information or reading), and consequently it is an area with a central visual field bias (Levy, Hasson et al., 2001; Malach, Levy, & Hasson, 2002; Hasson, Harel, Levy, & Malach, 2003). Our results however indicate that peripheral faces can also significantly activate the FFA, even when they are unattended (Figure 4.5c).

4.4.5 Saturating the BOLD signal—some thoughts

The main result of this study is that top-down behavioral relevance is sufficient for evoking high levels of BOLD activity in the FFA, even when the face stimuli are outside the focus of attention. However as we mentioned earlier, a possible concern with this interpretation could be that if the signal during the dual-task condition was saturated, we would necessarily not have observed any additional benefit from shifting attention onto the faces (i.e. the peripheral condition). We have already presented data (Section 4.3.2) that confirms that the voxels in our ROIs were not saturated, since for each of 4 subjects significantly higher levels of activity were observed during the localizer tasks with peripherally presented faces, compared to the dual-task experiment. However, these results do not necessarily preclude the possibility that the underlying neuronal populations were not saturated during the peripheral-task condition.

The primary argument against saturation in our analysis hinges on the supposition that presenting faces for longer durations would cause the underlying neuronal population to be more active, thereby increasing the overall signal measured in the relevant voxels. Indeed it is known that masked stimuli activate neuronal populations much less than unmasked stimuli presented for longer durations (Rolls & Tovee, 1994; Rolls, Tovee, & Panzeri, 1999; Rolls, 2004). This observation, however, is generally only true when neuronal activity is considered over the longer intervals of time during which the unmasked stimuli are present. At any particular moment in time when the stimulus is displayed, the firing rates for unmasked stimuli are not necessarily higher compared to those of masked stimuli. In other words, the optimal stimuli of the cell under study, whether masked or unmasked can saturate the neuronal response during the time they are presented. The reason the fMRI signal could possibly show differences in the levels of activity elicited by masked versus unmasked stimuli is that it integrates the underlying neuronal signal over a longer time scale, during which interval the effects of presenting stimuli for longer periods become apparent in the average neuronal activity.

For our purposes, from a neuronal point of view, the face stimuli could very well have saturated the relevant neural signal during the dual-task condition, at least during the time they were presented (before the onset of the mask). Thus at the level of the neurons responding to our peripheral stimuli, the additional effects of shifting the focus of attention onto the faces in the peripheral-task condition would not have been detected.

Does this possibility then undermine our main conclusions? Not really! Our conclusion is that, in the absence of focused attention, top-down behavioral relevance can produce high levels of neuronal activation. Whether saturation occurs or not this conclusion still holds—we see high levels of activity in the dual-task condition when focal attention is not on the faces. In fact, if saturation does occur, our results would imply that neuronal activities can reach their maximal level of response in the absence of focused attention! This result is indeed intriguing, given that current views hold that neuronal activity must be decreased in the absence of attention.

It must be noted that saturation could become a problematic issue if we wish to demonstrate that focal attention has absolutely no effect on BOLD activity—beyond that induced by behavioral relevance. But this is not our stance. Although no significant effect of focal attention on BOLD activity was observed in the particular conditions of our experiment, we are not opposed to the idea that spatial attention can increase the neuronal response under some conditions. Indeed when the stimulus is presented in a cluttered environment, or is degraded, several monkey electrophysiology and human fMRI studies show the pronounced effect of spatial attention, as we discuss below. However, our results suggest that, at least under some conditions, removing spatial attention does not necessarily result in lower levels of cortical activity. This finding, which holds independent of whether saturation occurs or not, contrasts with the previous understanding of face-related BOLD activity and focal attention.

4.4.6 Activity in other brain regions

This far, we have primarily focused on activity in the FFA because this region responds specifically to faces. However, it is of interest to also consider activity in other brain areas during the different task conditions. We localized the parahippocampal place area and the lateral occipital complex for our subjects, and our findings in these areas are discussed below.

4.4.6.1 The parahippocampal place area (PPA)

The parahippocampal place area is a region in the parahippocampal cortex that responds strongly to spatial layouts compared to faces and other objects (Epstein & Kanwisher, 1998). In our localizer task, we identified this area for all 12 of our subjects, and as expected from its known visual properties, the face and letter stimuli failed to elicit large levels of activation in the PPA (compared to the FFA; Figures 4.4c and 4.6c). In addition, we did not observe differential effects of focal attention and task relevance in these voxels. This is not entirely unexpected given the lack of a visual response in the first place. This finding is reassuring since, as discussed in section 4.3.2 it confirms that our observations are specific to the voxels we targeted and is not a whole-brain effect.

4.4.6.2 The lateral occipital complex (LOC)

The LOC, a cortical area known to be important in the processing of object shapes (Malach, Reppas et al., 1995; Kanwisher, Woods, Iacoboni, & Mazziotta, 1997; Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998; Kourtzi & Kanwisher, 2000, 2001), was localized for the six untrained subjects. Over these six subjects, we observed a significant increase in the LOC BOLD signal compared to baseline for both our central and peripheral tasks. This result is compatible with previous reports showing that letters and faces activate regions in the lateral occipital region (Puce, Allison, Asgari, Gore, & McCarthy, 1996). The activity elicited by the attended letter stimuli was significantly higher than that elicited by the attended faces (p=.02; Figure 4.8) suggesting that the ROI we localized in the lateral occipital cortex may exhibit a preference for letters over faces. However, we did not observe a significant effect of focal attention and behavioral relevance on the BOLD signal in the LOC in the different task conditions (p=.14, dual-task vs. peripheral-task condition).

4.4.6.3 Retinotopic visual areas

Although we did not localize the boundaries of the early visual areas in our subjects, it is worth speculating about what we would have observed in retinotopically organized regions. Several fMRI studies have reported that focal attention modulates activity in early areas including V1 (Tootell, Hadjikhani et al., 1998; Watanabe, Harner et al., 1998; Sengpiel & Hubener, 1999; Somers, Dale et al., 1999), and thus it is likely that different patterns of activity would have been observed in response to the peripheral and central stimuli in the different task conditions. In particular, it is probable that activity at foveal representations would have been higher in the central and dual-task conditions (compared to the peripheral-task conditions) since focal attention would have been directed to the foveal stimuli. Conversely, at the peripheral representations, activity would have been expected to be lower during the central task condition compared to the peripheral task condition. But it is interesting to consider what would happen at peripheral locations during the dual-task condition, when focal attention is not available to the periphery, but the peripheral stimuli are behaviorally relevant. Because we observe a modulation of activity due to top-down behavioral relevance in higher areas (FFA), it is legitimate to ask whether the same effect arises in earlier retinotopic areas. This is unlikely however since these retinotopic areas do not explicitly encode our peripheral stimuli (faces). The fact that these effects are not observed in the PPA, which also gets input from the early visual cortex, probably implies that the effect is not present in such retinotopic areas. This does not rule out the possibility however that top-down behavioral relevance could affect activity in these areas when appropriate stimuli (such as sinusoidal gratings) are used.

4.4.7 In comparison with monkey electrophysiology

With reference to electrophysiology studies, our results showing the absence of a significant effect of focal attention on BOLD activity might seem to contradict previous reports of strong attentional modulation at the neuronal level (Moran & Desimone, 1985; Spitzer, Desimone et al., 1988; Reynolds, Chelazzi et al., 1999; Reynolds, Pasternak et al., 2000; Reynolds & Desimone, 2003). However, it is important to note that these observations have generally been made when two competing stimuli (a preferred and a non-preferred stimulus) are placed within the receptive field (RF) of a recorded neuron. When a single stimulus is placed within the RF (in the absence of competition), the effect of shifting attention in and out of the RF is much less pronounced, at least for well-contrasted stimuli (Moran & Desimone, 1985; Haenny, Maunsell et al., 1988; Maunsell, Sclar et al., 1991; Motter, 1993; Luck, Chelazzi et al., 1997; Reynolds, Chelazzi et al., 1999). Attentional effects compatible with these single-cell observations have also been reported in fMRI studies (Kastner, De Weerd et al., 1998). Thus, far from being inconsistent, electrophysiological accounts are compatible with the data we report here—namely that, for well-isolated and contrasted stimuli (such as our face stimuli), the effect of removing the focus of attention alone, while keeping task demands comparable, does not result in a substantial drop in neuronal activity.

4.5 Conclusions

In conclusion, our results might help explain a puzzling discrepancy observed at the behavioral level: in some cases, visual processing can fail dramatically in the absence of attention (e.g. change-blindness or inattentional blindness (Rock, Linnett, Grant, & Mack, 1992; Simons & Levin, 1997)) while in other cases (dual-task), perception outside the focus of attention can be quite successful. When visual stimuli or changes are totally unexpected, absence of top-down behavioral relevance might decrease their associated neuronal representations, as shown here. On the other hand, as soon as the range of possible relevant events can be constrained by top-down influence (as in dual-task), neural activities need not suffer even in the absence of focal attention.

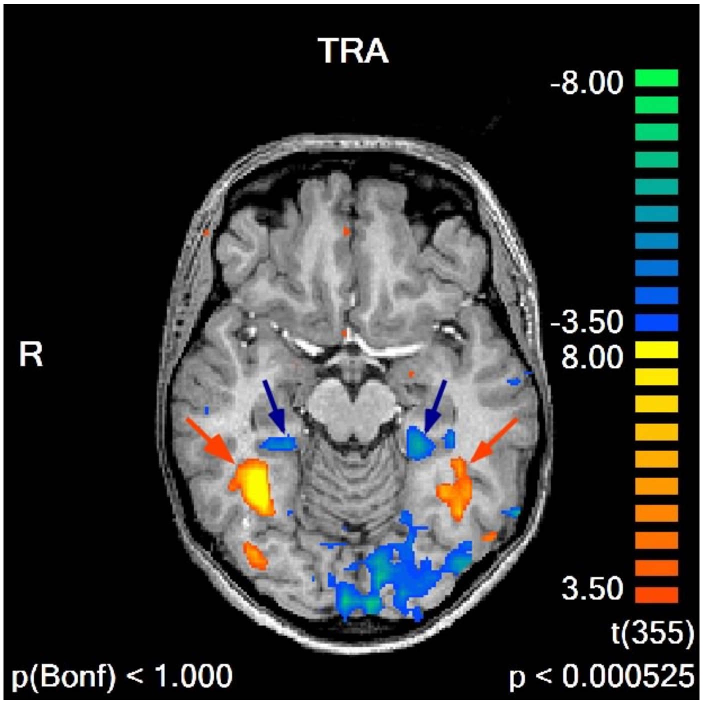

Figure 4.1: Localizing the FFA and PPA. The FFA (orange arrows) and PPA (blue arrows) localized for one of our subjects are shown in this transverse slice. The FFA and PPA were localized by contrasting BOLD activity during blocked presentation of color faces against blocked presentation of outdoor scenes. The mean Talairach coordinates of the FFA over all subjects were: Right FFA: x = 38 ± 1mm, y = -50 ± 2 mm, z = -13 ± 1mm; Left FFA: x = -40 ± 1mm, y = -51 ± 2mm, z = -15 ± 2mm. The mean Talairach coordinates of the PPA over all subjects were: right PPA: x = 24 ± 2mm, y = -42 ± 2mm, z = -6 ± 1mm; left PPA: x = -28 ± 1mm, y = -46 ± 4mm, z = -9 ± 2mm.

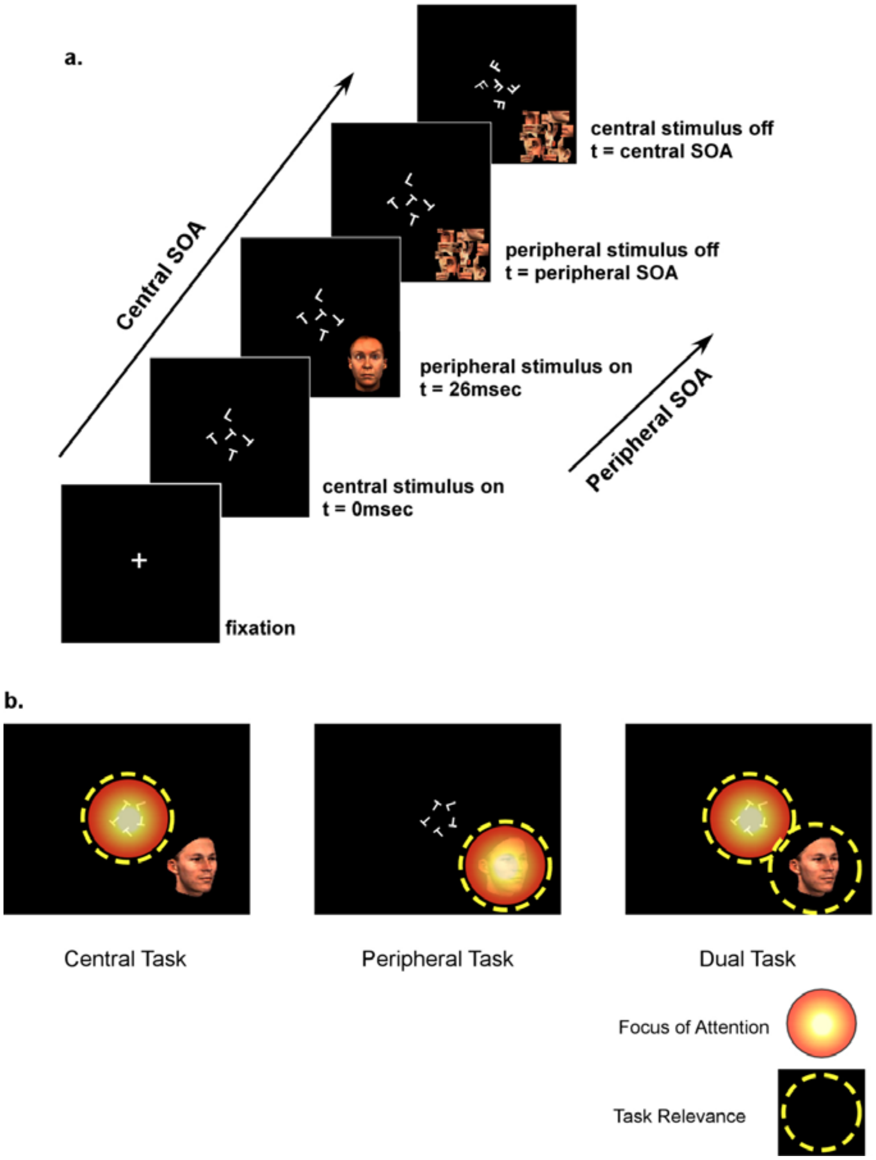

Figure 4.2: Manipulating attention and behavioral relevance in the dual-task experiment. a) Schematic timeline for a typical trial in the dual-task. At the end of the trial, subjects are required to report the gender of the face presented and/or whether the letters were all the same or different. The layout was the same for all trials, independent of specific instructions given to subjects at the beginning of each block. Central SOA(~200 ms) and peripheral SOA (~160 ms) indicate the presentation time for letters and faces, respectively. b) A simplified explanation of the manipulation of the focus of attention and behavioral relevance in the 3 conditions of the paradigm. In the central-task condition, the focus of attention is on the letters, and since subjects report on the letters, the faces are behaviorally irrelevant. In the peripheral-task condition, the focus of attention is available to the faces, and the faces are also behaviorally relevant. However, in the dual-task condition, the focus of attention is away from the faces, but they are still behaviorally relevant since subjects make a report on them. Thus, in this paradigm, we have conditions in which the faces are relevant and attended, a condition in which we remove only the focus of attention, and finally a condition in which we also make the faces behaviorally irrelevant. The three conditions enable us to tease apart the effects of these two factors.

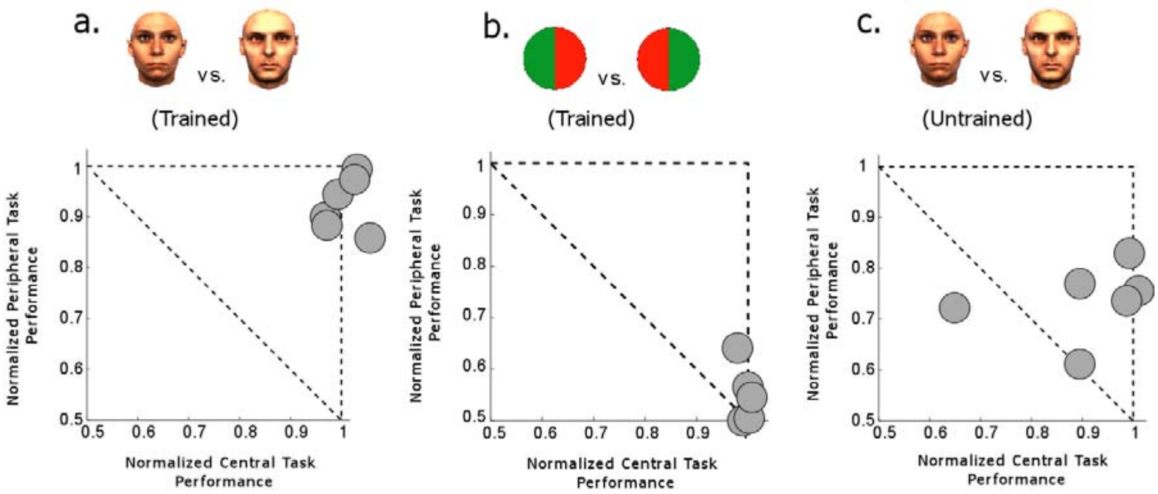

Figure 4.3: Behavioral performance. Dual-task performance values relative to performance achieved in the single-task conditions. a) Face-gender discrimination performance for 6 trained subjects obtained in the scanner. Each point represents a subject's dual-task performance on the central and peripheral tasks relative to their single-task performance. This group of subjects, achieves a high level of performance (average performance in the dual-task condition: 92.5% of performance in the single-task condition). The performance of these subjects and those reported in Chapter 2 were comparable (performance on letter discrimination task, p=0.79; performance on face-gender task, p=0.69). b) Performance for the same 6 subjects on a known attentionally demanding disk discrimination task (outside the scanner) falls to chance levels in the dual-task condition. Thus, the central task is effective in withdrawing focal attention away from the periphery. c) Face-gender performance for 6 new untrained subjects (in the scanner). As expected, we observe a significant drop (p<.009) in performance for these subjects in the dual-task condition.

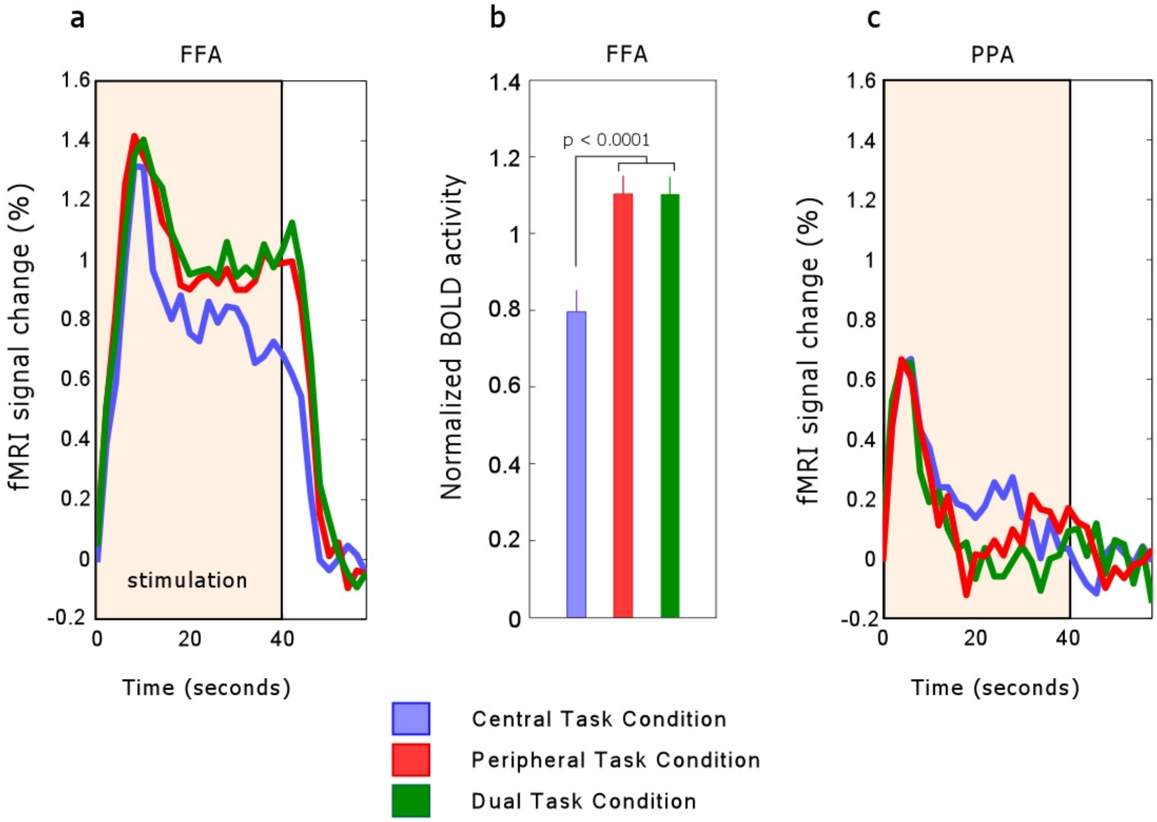

Figure 4.4: BOLD activity for 6 trained subjects in the dual-task condition. a) Time course of FFA activity. The data shows a significant drop in activity in the central task condition compared to the peripheral task condition. However, there is no significant drop in activity in the dual-task condition. Thus, removing the focus of attention alone does not affect activity. Making the faces task-irrelevant results in a significant decrease in the signal. b) The same data shown in a different format for the 3 conditions. Each bar corresponds to the area under the 3 curves shown in a) over the 40 s stimulation period (shifted by 7 s to take into account the delay in the hemodynamic response). A value of 1 corresponds to the average activation in the ROIs over all stimulation periods of the entire experiment. c) Activity in the PPA. The trend observed in the FFA is not observed in other brain areas, thus ruling out the possibility that the observed effects are due to general arousal.

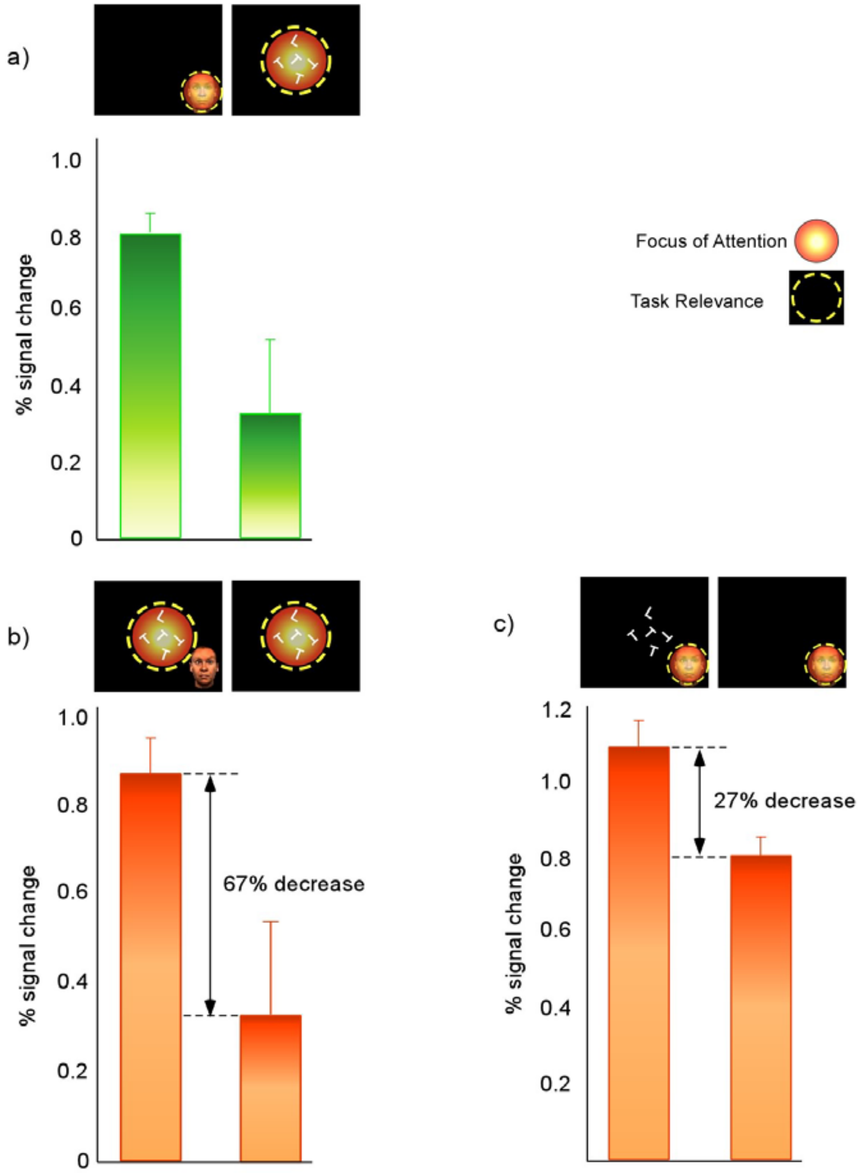

Figure 4.5: Estimating the respective contribution of letters and faces to the observed BOLD signal in the FFA. BOLD activity in the FFA for 3 subjects from the main experiment (Figure 2) and a control experiment in which they were presented with either only letters, or only faces, and performed the corresponding discrimination tasks. a) In the complete absence of the other stimulus, attended (and behaviorally relevant) faces activate the FFA 2.7 times more than attended (and behaviorally relevant) letters. These voxels are thus considerably more responsive to faces. b) The effect of removing the unattended face. The data on the left is obtained from the central task condition for these 3 subjects when the letters were attended but the faces were also presented. In the second bar, data from the control experiment is shown in the condition in which only letters were presented (and attended to) (note this is the same data as on the right in a)). Removing the faces from the display results in a large (67%) decrease in the FFA BOLD signal. c) In contrast, removing unattended letters while faces are presented results in only a 27% decrease. The data on the left is obtained from the peripheral task condition. On the right is data from the control experiment when the faces were attended and the letters were absent (left bar in a).

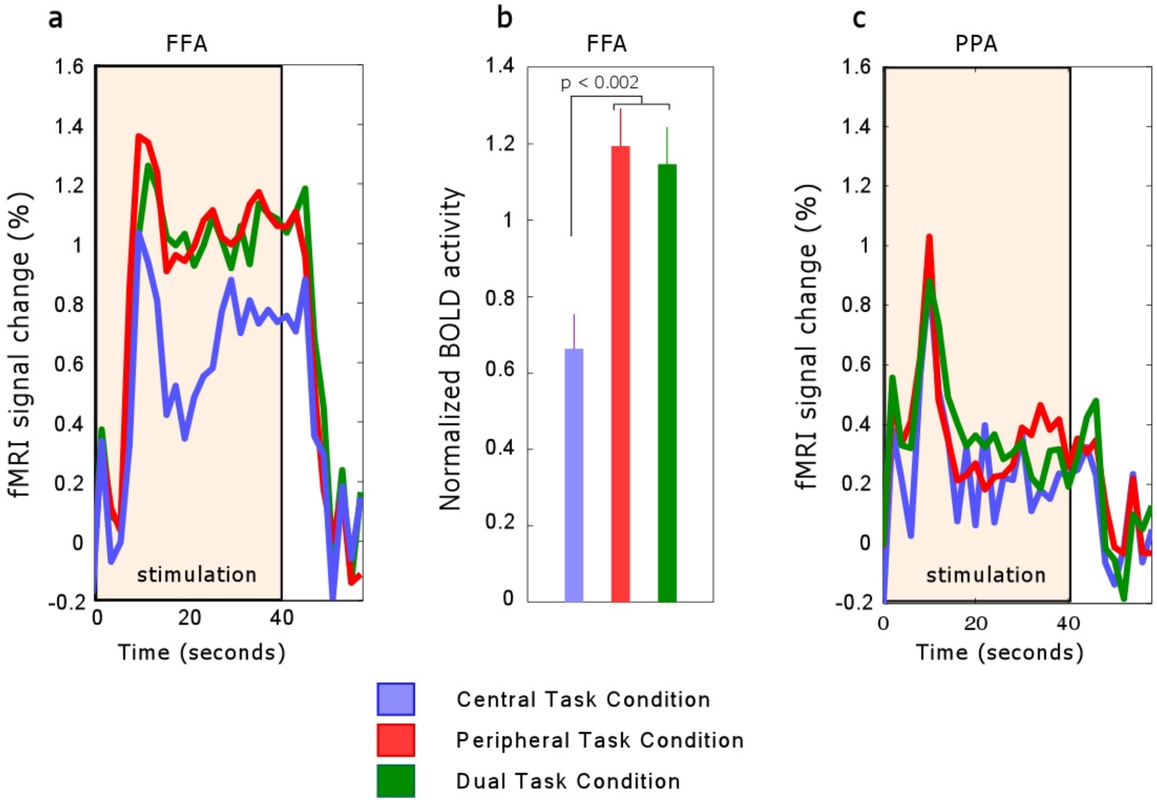

Figure 4.6: BOLD activity for 6 untrained subjects in the dual-task condition. a) Time course of FFA activity. Similar to Figure 3, removing the focus of attention alone does not affect BOLD activity. Instead, making the faces task-irrelevant results in a significant decrease in the signal. b) The same data shown in a different format for the 3 conditions. Note that despite the substantial drop in behavioral performance in the dual-task condition (Fig 4.3c), no corresponding decrease in BOLD activity is observed in this condition. c) BOLD activity in the PPA. The trend observed in the FFA is not observed in another brain area for the untrained subjects, thus ruling out the possibility that the observed effects are due to general arousal.

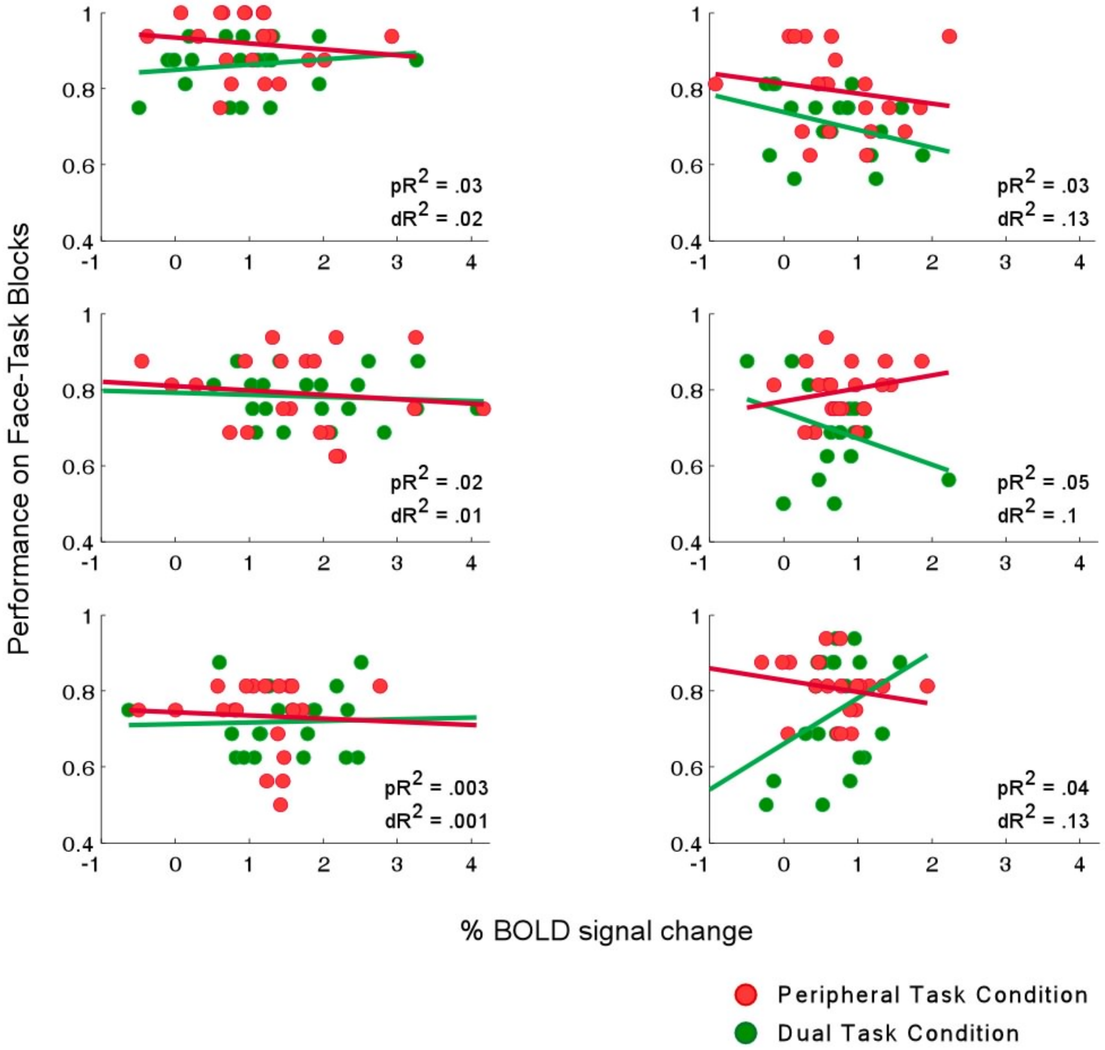

Figure 4.7 Correlating FFA BOLD activity with behavioral performance. The data shown here is for the six ‘trained’ subjects. For each subject performance on each face task block (i.e., peripheral and dual blocks) is plotted against the % BOLD signal change measured on that block. The red circles and lines represent the peripheral task condition, and the green circles and lines represent the dual task condition. The correlation coefficients for the peripheral (pR^2) and dual-task conditions (dR^2) for the six subjects are indicated on each graph. On average, over the trained and untrained subjects the correlations were weak (r^2<0.13).

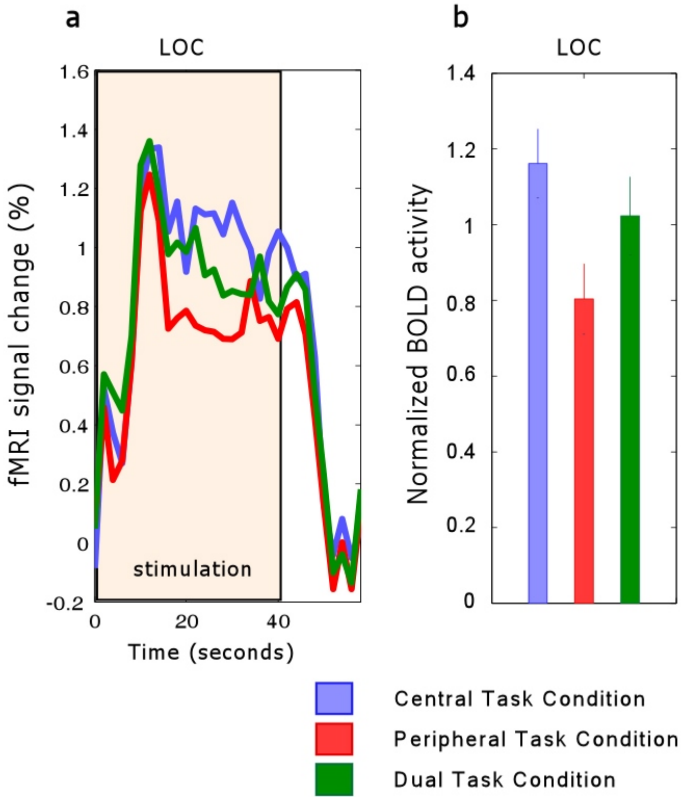

Figure 4.8: BOLD activity for the untrained subjects in the dual-task condition in the LOC. a) Time course of LOC activity. In the LOC, in contrast to the FFA, the opposite trend is observed, namely that BOLD activity in the central-task condition is higher than activity in the peripheral task condition (p=.02), indicating that the attended letters activate these voxels more than attended faces do. The difference in activity between the dual- and peripheral-task conditions is not significant (p=.14). b) The same data shown in a different format for the 3 conditions.