

Activity in human V1 follows multisensory perception

S. Watkins,^{a,b,*} L. Shams,^c O. Josephs,^a and G. Rees^{a,b}

^aWellcome Trust Centre for Neuroimaging, University College London, 12 Queen Square, London WC1N 3BG, UK

^bInstitute of Cognitive Neuroscience, University College London, Alexandra House, 17 Queen Square, London WC1N 3AR, UK

^cDepartment of Psychology, University of California, Los Angeles, Los Angeles, CA 90095-1563, USA

Received 5 December 2006; revised 11 April 2007; accepted 3 May 2007
Available online 25 May 2007

When a single brief visual flash is accompanied by two auditory beeps, it is frequently perceived incorrectly as two flashes. Such illusory multisensory perception is associated with increased activation of retinotopic human primary visual cortex (V1) suggesting that such activity reflects subjective perception [Watkins, S., Shams, L., Tanaka, S., Haynes, J.D., Rees, G., 2006. Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage*, 31, 1247–1256]. However, an alternate possibility is that increased V1 activity reflects either fluctuating attention or auditory–visual perceptual matching on illusion trials. Here, we rule out these possibilities by studying the complementary illusion, where a double flash is accompanied by a single beep and perceived incorrectly as a single flash. We replicate findings of increased activity in retinotopic V1 when a single flash is perceived incorrectly as two flashes, and now show that activity is decreased in retinotopic V1 when a double flash is perceived incorrectly as a single flash. Our findings provide strong support for the notion that human V1 activity reflects subjective perception in these multisensory illusions.

© 2007 Elsevier Inc. All rights reserved.

Introduction

In everyday life our perception of the world is dominated by multisensory information. Multisensory convergence can influence not only cortical sensory processing (for a review see Foxe and Schroeder, 2005) but also the consciously perceived properties of stimuli (McGurk and MacDonald, 1976; Mottonen et al., 2002; Murray et al., 2004, 2005; Shams et al., 2000; Stein et al., 1996). However, there has been relatively little study of how changes in conscious perception associated with multisensory interactions might be reflected in changes in brain activity. Recently, we used high field fMRI to study brain activity associated with an

established audiovisual illusion. When a single brief visual flash is accompanied by two auditory beeps, it is frequently perceived incorrectly as two flashes (Shams et al., 2000). We found that perception of this ‘fission’ illusion is associated with increased activity in retinotopic areas of human primary visual cortex representing the visual stimulus (Watkins et al., 2006).

Such an association of V1 activity with illusory multisensory perception is consistent with earlier findings that visual-evoked potentials and fields are modified at short latency in association with the illusion (Bhattacharya et al., 2002; Shams et al., 2001, 2005). Moreover, it may suggest that activity in V1 reflects subjective perception rather than the visual stimulus that was physically presented. However, an alternate possibility is that enhanced V1 activity for the multisensory ‘fission’ illusion might represent the effects of either fluctuating attention, or a non-specific response to a perceptual matching between sensory modalities, rather than a response that truly varied with perception.

Here, we sought to rule out these possibilities by replicating our earlier findings and now comparing them with a complementary illusion (Andersen et al., 2004). In contrast to the previous work, which focused on situations where one physical flash was incorrectly perceived as two flashes (‘fission’), here we now focused on situations where two physical flashes are incorrectly perceived as one flash (‘fusion’). If V1 activity reflects subjective perception, then it should be enhanced for the ‘fission’ but reduced for the ‘fusion’ illusion, reflecting the illusory perception of two (when one was physically present) or one (when two were physically present) flashes respectively. However, an account of the illusion that postulates V1 responses reflecting the modulatory influences of attention or auditory–visual matching predicts that V1 activity should be enhanced both for ‘fission’ and ‘fusion’ illusions.

Methods

Subjects

Fourteen young adults (6 females, 18–30 years old, right handed) with normal hearing and normal or corrected to normal

* Corresponding author. Institute of Cognitive Neuroscience, University College London, Alexandra House, 17 Queen Square, London WC1N 3AR, UK. Fax: +44 20 7813 1420.

E-mail address: swatkins@fil.ion.ucl.ac.uk (S. Watkins).

Available online on ScienceDirect (www.sciencedirect.com).

vision gave written informed consent to participate in the study, which was approved by the local ethics committee. Prior to scanning all subjects took part in a behavioral pilot experiment (see procedure for full details), following which two subjects were excluded because they did not report the multisensory illusion. Following scanning, two subjects were rejected on the basis of excessive head movement (>5 mm). Ten subjects (6 females, 18–30 years old, right handed) were therefore included in the analysis reported here.

Stimuli

Visual stimuli were projected from an LCD projector (NEC LT158, refresh rate 60 Hz) onto a circular projection screen at the rear of the scanner. The subjects viewed the screen via a mirror positioned within the head coil. The auditory stimuli were presented binaurally using electrostatic headphones (KOSS, Milwaukee, USA, Model: ESP 950 Medical) custom adapted for use in the scanner. All stimuli were presented using MATLAB (Mathworks Inc.) and COGENT 2000 toolbox (www.vislab.ucl.ac.uk/cogent/index.html). Visual stimuli consisted of an annulus with luminance 420 cd/m^2 and eccentricity $8.5\text{--}10^\circ$ of visual angle presented for 17 ms. When two flashes were presented, the interval between them (ISI) was 46 ms. The background was a uniform grey screen of luminance 30 cd/m^2 . Luminance calibration was achieved via a viewing aperture in the MRI control room using a Minolta LS-100 spot photometer. We used an annulus displayed in the peripheral visual field in association with auditory stimulation to maximise illusory perception, which is stronger for stimuli displayed in the periphery (Shams et al., 2002). In addition, the cortical representation of such a peripheral annulus avoids the foveal confluence at the occipital pole (Serenio et al., 1995), where it is extremely difficult to distinguish activity from different early retinotopic visual cortical areas. Our stimulus geometry therefore permitted us to clearly distinguish activity in V1, V2 and V3 from other cortical areas. The auditory stimuli consisted of a sine wave with frequency 3.5 kHz, duration 10 ms (with a ramp time of 1 ms at each end of the sound wave envelope) and volume 95 dB. The sound intensity (SPL) produced by the headphones was measured while the headphones were a suitable distance away from the scanner using a sound meter (Radioshack 33-2055). When two bleeps were presented, the interval between them (ISI) was 46 ms. On trials with two flashes and one bleep, the auditory bleep was presented simultaneously with the first flash. Similarly, on trials with two bleeps and one flash the flash was presented simultaneously with the first bleep. Pilot behavioral work confirmed previous observations that whether bleeps and flashes are presented simultaneously or with slight temporal offset (Shams et al., 2002) makes little difference to behavioral reports of illusory perception.

Procedure

Subjects initially took part in a behavioral pilot study. Subjects were presented with one or two briefly and successively flashed visual stimuli, either alone or accompanied by one or two successively presented auditory bleeps. For clarity, these trial types will be referred to by abbreviations. For example, ‘F2B1’ refers to trials on which there were two flashes and one bleep while ‘F2B2’ refers to a trial on which two flashes and two bleeps were presented. Subjects were instructed to report by button press

whether they perceived one or two flashes and ignore the bleeps. Each participant completed 1 run of 128 trials divided between the different trial types (F1B1, F1B2, F2B1, F2B2, F1 and F2) while in the scanner. This pilot study ensured that the subjects experienced the multisensory illusion and could clearly distinguish one and two flashes with no auditory stimulation (all subjects were able to achieve $>95\%$ correct on visual alone trials before starting the main experiment). The two subjects who were excluded showed no difference in error rates in the F1B1 and F2B2 conditions (5% and 4% error rates for each excluded subject), compared to 6% (S.D. 1%) for the group who went forward to the experiment.

During the fMRI experiment, on each trial subjects were presented with one or two briefly and successively flashed visual stimuli accompanied by one or two successively presented auditory bleeps. These comprised four different trial types that represented all the possible combinations of flashes and bleeps. Subjects maintained central fixation throughout and indicated whether they perceived one or two flashes, by pressing one of two response keys on a keypad held in their right hand. Each trial lasted 90 ms followed by a 1800 ms response interval. Eye position data was collected from all subjects during the trials to ensure subjects maintained fixation. One seventh of all trials were null trials, during which no visual or auditory stimuli were presented. There were thus five physically different types of trial. The responses of subjects were further used to post hoc divide the F2B1 and the F1B2 trials into those on which the illusion was perceived (“F2B1-Fusion Illusion” and “F1B2-Fission Illusion”), and those on which it was not (“F2B1-no Illusion” and “F1B2-no Illusion”). Each subject completed between 4 and 6 runs of 112 trials divided equally between the different trial types. Trials were pseudo-randomly distributed within a run.

fMRI scanning

A 3T Siemens Allegra system was used to acquire both T2*-weighted echoplanar (EPI) images with Blood Oxygenation Level Dependent contrast (BOLD) and T1 weighted anatomical images. Each EPI image comprised of thirty-two 3 mm axial slices with an in-plane resolution of 3×3 mm positioned to cover the whole brain. Data were acquired in four to six runs, each run consisting of 162 volumes. The first five volumes of each run were discarded to allow for T1 equilibration effects. Volumes were acquired continuously with a TR of 2.08 s per volume. During scanning, eye position and pupil diameter were continually sampled at 60 Hz using long-range infrared video-oculography (ASL 504LRO Eye Tracking System, Mass). Eye movements were monitored on-line via a video screen for all subjects. Subjects completed a short pilot in the scanner to ensure they could maintain fixation.

Data analysis

Eye tracking data were analysed with MATLAB (Mathworks Inc.). Blinks and periods of signal loss were removed from the eye movement data. Mean eye position, expressed as a distance from fixation, was then computed for each trial type and every subject from whom data were available. A repeated-measures ANOVA was used to establish whether mean eye position deviated significantly from fixation, or between conditions.

fMRI preprocessing

The EPI magnitude images undergoing statistical analysis were reconstructed from the complex k -space raw data using a generalized reconstruction method based on the measured EPI k -space trajectory to minimize ghosting (Josephs et al., 2000). Prior to reconstruction the k -space raw data were assessed for spike artefacts as indicated by high background noise (two-fold oversampling in the readout direction always allowed for estimating the background noise from areas outside the head) (Weiskopf et al., in press). If k -space phase-encoding lines were affected by spikes they were replaced by the corresponding k -space lines from adjacent uncorrupted time points of the EPI time series. A correction for linear phase variations across k -space (due to inter-scan motion) was applied prior to replacing the data. Replacing single k -space lines instead of complete slices or volumes ensured that a minimal amount of data was interpolated. Less than 0.03% of all k -space lines required correction, thus minimally affecting the experimental degrees of freedom. The spike detection and correction were implemented in MATLAB (Mathworks Inc.).

The resulting functional imaging data were analyzed using Statistical Parametric Mapping software (SPM2, Wellcome Department of Imaging Neuroscience, University College London). All image volumes were realigned spatially to the first, and temporally corrected for slice acquisition time (using the middle slice as a reference). Resulting image volumes were coregistered to each subject's structural scan. The fMRI data were analyzed using an event-related model. Activated voxels in each experimental condition for each subject were identified using a statistical model containing boxcar waveforms representing each of the experimental conditions, convolved with a canonical hemodynamic response function and mean corrected. Motion parameters defined by the realignment procedure were added to the model as six separate regressors of no interest. Multiple linear regression was then used to generate parameter estimates for each regressor at every voxel for every run. The resulting parameter estimates were averaged across runs to give a final parameter estimate for each of the experimental conditions for every subject. In order to get an accurate parameter estimate for each condition, any run with less than 4 events in a given illusion condition was excluded from the analysis. Data were scaled to the global mean of the time series and high pass filtered (cut-off: 0.0083 Hz) to remove low-frequency signal drifts.

Retinotopic analyses

To identify the boundaries of primary visual cortex, standard retinotopic mapping procedures were employed (Serenio et al., 1995; Teo et al., 1997; Wandell et al., 2000). Only 7 of the 10 subjects participated in the retinotopic mapping procedures and so data from these 7 are reported here, while the non-retinotopic analyses reported below used all 10 subjects. There were no behavioral or demographic differences between the two groups. Flashing checkerboard patterns covering either the horizontal or vertical meridian were alternated with rest periods for 16 epochs of 26 s over a scanning run lasting 165 volumes. SPM2 was used to generate activation maps for the horizontal and vertical meridians. Mask volumes for each region of interest (left and right V1, V2, and V3) were obtained by delineating the borders between visual areas using activation patterns from the meridian localisers. We

followed standard definitions of V1 together with segmentation and cortical flattening in MrGray (Teo et al., 1997; Wandell et al., 2000). Using the mask volumes for left and right V1, V2 and V3, we identified voxels that showed significant activation ($p < .05$ uncorrected) for the comparison of all trials on which visual stimulation was present (i.e. all experimental conditions) compared to null events, employing the regression analysis described above. This comparison identifies voxels activated by the annular visual stimulus in each of the retinotopic areas. Informal examination of these activations superimposed on flattened representations of occipital cortex confirmed our expectation that they represented voxels activated by our annular visual stimulus (Fig. 1).

Having thus independently identified the stimulus representation in V1–V3, we then extracted and averaged the regression parameters from the analysis of the main experimental time-series (described above). This procedure yielded estimates of percentage signal change for each condition averaged across voxels in V1, V2 and V3 that responded to the visual stimulus. The statistical significance of any differences in activation between the Illusion condition and the No-Illusion conditions was assessed by entering the percentage signal change for each subject in each condition into a two-tailed t -test using a significance level of $p < .05$. Finally, we calculated a mask representing the voxels in V1 that did not show a significant response to the visual stimulus. We then used this image to repeat the above procedure to examine the response to each condition in the non stimulus responsive area of V1.

Whole brain analysis

To complement the retinotopic analyses, we also conducted an unbiased examination of regions outside retinotopic cortex using a random-effects whole-brain analysis of all 10 subjects. The realigned and slice time corrected images from each subject were spatially normalized to a standard EPI template volume based on the MNI reference brain in the space of Talairach Tournoux (1988). The

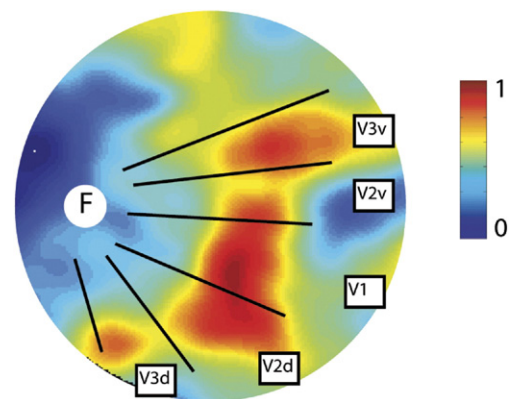


Fig. 1. Spatial distribution of stimulus-evoked activity in retinotopic visual cortex. The statistical contrast of all visual events (F1B1, F1B2, F2B1, F2B2) versus null events thresholded at $p < .05$ uncorrected is shown projected onto a flattened representation of visual cortex for a representative subject (the letter 'F' represents the location of the fovea, corresponding to the occipital pole; and the colorscale represents the t value at each location for the statistical contrast above, where red represents highest t values and blue the lowest; see Methods for full details). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

normalized image volumes were then smoothed with an isotropic 9 mm FWHM Gaussian kernel. These data were analyzed using an event-related random-effects model, the first stage of which was identical to the regression model described above for the retinotopic analyses, except now applied to spatially normalised images. The parameter estimates for different conditions were then entered into a second level analysis using planned comparisons with paired *t*-tests. For these whole brain analyses, a statistical threshold of $p < .05$ corrected for multiple comparisons was used except for areas previously associated with the fission illusion where a small volume correction (sphere of diameter 3 mm centered on coordinates [54 –54 30]) was applied (Watkins et al., 2006).

Results

Behavioral

Analysis of behavioral responses during scanning confirmed that subjects were able to accurately report the number of flashes when the number of flashes and bleeps were identical (i.e. F1B1 and F2B2 trial types; accuracy 94%, SE across subjects 1%). On a large proportion of trials when two flashes were accompanied by one bleep (F2B1 trials), subjects reported an illusory perception of one flash (“F2B1-Fusion Illusion”; 42% of all F2B1 trials, SE across subjects 6%). On the remainder of F2B1 trials, subjects reported veridical perception of two flashes (“F2B1-no Illusion”). When one flash was accompanied by two bleeps (F1B2) subjects reported an illusory perception of two flashes (F1B2-Fission illusion) on 34% of the trials (SE across subjects 7%). The frequency of occurrence of the fusion illusion was not significantly different from the Fission illusion (42% vs 44% $t(9)=1.24$, $p=.25$).

Eye position data

Subjects were requested to maintain fixation at the center of the display. During scanning eye position was monitored on-line in all subjects to ensure subjects successfully maintained fixation throughout the experiment sessions. A repeated-measures ANOVA showed no statistical difference in mean eye position from fixation, or between conditions for the eight subjects in whom eye data were available ($F(6,42)=.957$, $p=.466$). Eye data was monitored but not recorded in two subjects.

Functional MRI

Retinotopic analyses

Many F2B1 trials (42%) evoked the illusion of one flash (F2B1-Fusion Illusion), while on the remainder two flashes were perceived (F2B1-no Illusion). We therefore compared activity in retinotopic visual areas that was evoked on F2B1-Fusion Illusion trials with F2B1-no Illusion trials and on F1B2-Fission Illusion trials with F1B2-no Illusion trials in the seven subjects where retinotopic maps were obtained. Stimulus-evoked activity in V1 was significantly lower for F2B1-Fusion Illusion trials on which the illusion was perceived, compared to F2B1-no Illusion when the illusion was not perceived ($t(6)=2.93$, $p=.026$, two-tailed) (see Fig. 2 for full details, including time courses). Note that we compared physically identical F2B1 trials with exactly the same visual and auditory stimulation that resulted either in the fission illusion or no illusion. Thus, any differences in brain activity

associated with this comparison *cannot* reflect differences in visual or auditory stimulation. Stimulus-evoked activity in V1 in the F2B1-Fusion Illusion condition (where one flash was perceived) was not significantly different from the F1B1 condition (where one flash was physically present) [$t(6)=.79$; $p=.45$] (see Fig. 2 for full details, including time courses).

Similarly, many F1B2 trials (34%) led to the illusion of two flashes (F1B2-Fission Illusion), while on the remainder only one flash was perceived (F1B2-no Illusion). We found that stimulus-evoked activity in V1 was significantly higher for F1B2-Fission Illusion trials on which the illusion was perceived compared to F1B2-no Illusion when the illusion was not perceived ($t(6)=2.70$, $p=.035$, two-tailed) (see Fig. 2). The activity in V1 in the F1B2-Illusion condition (where two flashes were perceived) was not significantly different from the F2B2 condition (where two flashes were physically present) [$t(6)=.32$; $p=.75$]. This replicates our previous findings (Watkins et al., 2005). Again, because physically identical trials are compared these differences cannot be attributed to differences in sensory stimulation. The activity in stimulus-driven regions of V2 and V3 showed a similar pattern of activation to V1 but did not reach conventional levels of statistical significance (Supplementary Fig. 1).

These differential cortical responses to the fission and fusion illusory perception were specific to the retinotopic locations of V1 responding to the visual annulus, as there was no significant effect of the illusions in the regions of V1 that did not respond to the visual annulus [Fission: $t(6)=.80$, $p=.46$ Fusion: $t(6)=.83$, $p=.45$]. Similarly, there was no evidence for a general effect of either judgment or number of flashes outside regions of V1 responsive to the visual annulus. Specifically, there was no significant difference between activity evoked in locations of V1 that did not correspond to the visual annulus for the judgment of 1 flash (i.e. conditions F1B1, F1B2-no Illusion and F2B1-Fusion Illusion) versus 2 flashes (i.e. conditions F2B2, F2B1-no Illusion and F1B2-Fission Illusion) or the actual presence of 1 flash (i.e. conditions F1B1 and F1B2) versus 2 flashes (i.e. conditions F2B1 and F2B2) [$t(6)=.21$, $p=.91$, $t(6)=2.1$, $p=.08$]. Finally, there were no significant differences between the F1B1 and F2B2 conditions in the non-stimulus responsive area of V1 [$t(6)=1.48$, $p=.19$].

Whole brain analyses

To complement the retinotopic analyses, we also performed whole-brain analyses of activity for each of the main comparisons outlined above. Unrestricted whole-brain analysis of illusory multisensory fusion perception (i.e. F2B1-Fusion Illusion vs F2B1-no Illusion) revealed significant activation in the right superior temporal sulcus ([58 –32 20]; $t=7.30$; $p=.01$ corrected at cluster level, number of voxels in the cluster=96). These activated loci are shown in Fig. 3. There were no cortical areas that showed a significant response to F2B1-no Illusion > F2B1 Fusion Illusion.

Unrestricted whole-brain analysis of illusory multisensory fission illusion (i.e. F1B2-Fission Illusion vs F1B2-no Illusion) revealed no significant cortical activation outside early visual areas at a corrected threshold. An examination of cortical areas previously associated with this illusion revealed significant activation in the right superior temporal sulcus ([52 –54 28]; $t=2.9$; $p=.04$, corrected for small volume examined).

Discussion

Our behavioral findings demonstrated that subjects perceived an illusory perception of one flash (“fusion”) rather than the

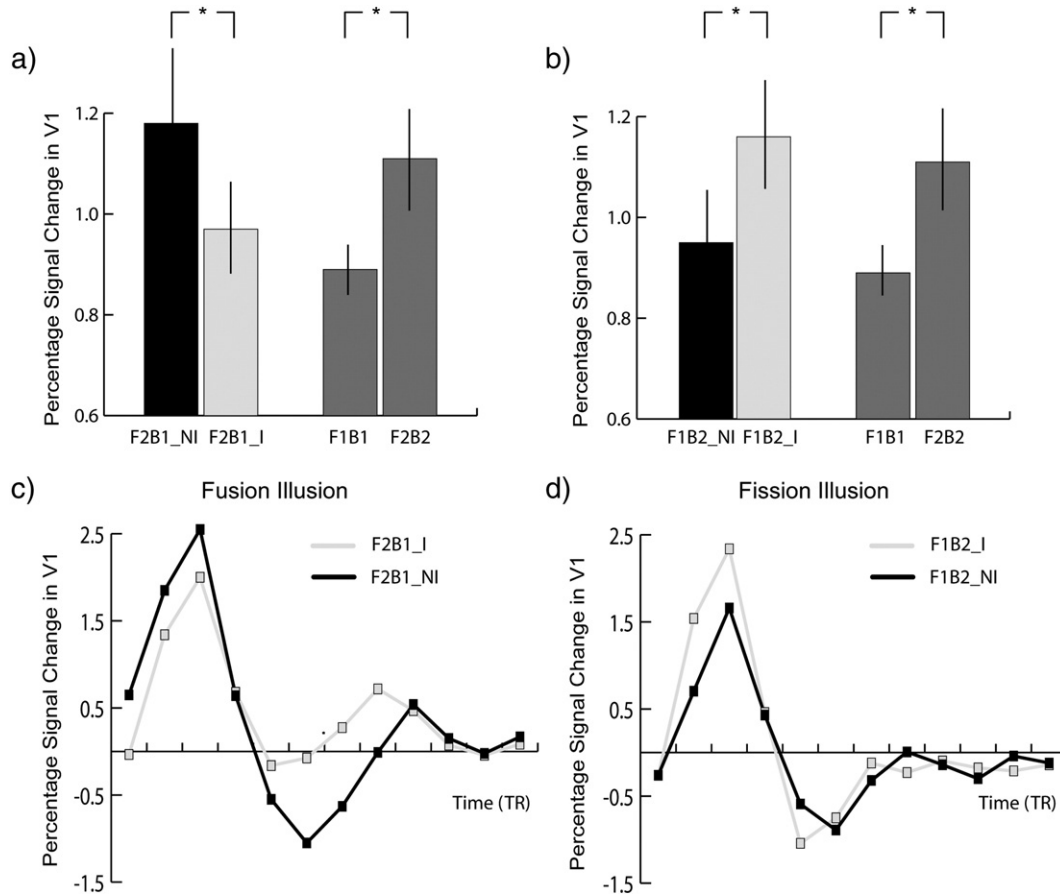


Fig. 2. Signal change in primary visual cortex associated with illusory multisensory perception. (a) The mean percentage signal change in retinotopically defined V1 (see Methods) is shown for the condition F2B1-no Illusion (two flashes with one beep when subjects reported correctly the perception of two flashes), F2B1-Fusion Illusion (two flashes with one beep when subjects reported the illusory perception of one flash), F1B1 and F2B2. (b) The mean percentage signal change in retinotopically defined V1 is shown for the condition F1B2-no Illusion (one flash with two beeps when subjects reported correctly the perception of one flash), F1B2-Fission Illusion (one flash with two beeps when subjects reported the illusory perception of two flashes), F1B1 and F2B2. Data shown are averaged across the seven subjects (see Methods for further details) with error bars representing the standard error of the mean, and the symbol “*” indicating statistical significance ($p < .05$). (c) Time courses for the V1 cortical responses in the F2B1-Fusion Illusion (grey line) and F2B1-no Illusion (black line) conditions. (d) Time courses for the V1 cortical responses in the F1B2-Fission Illusion (grey line) and F1B2-no Illusion (black line) conditions. Percentage signal change in V1 is plotted against time from stimulus onset (units of TR=2.08 s) for both conditions averaged across subjects. The time courses were calculated for each of the subjects by using a statistical model containing a boxcar waveform representing each of the experimental conditions, convolved with a series of FIR (finite impulse response) functions (using the SPM toolbox MarsBaR; <http://marsbar.sourceforge.net/>). Motion parameters defined by the realignment procedure were added to the model as six separate regressors of no interest. Multiple linear regression was then used to generate parameter estimates for each regressor at each time point for every subject. The data used in this model were extracted from the area of V1 that responded to the visual stimulus. This was determined by masking the entire V1 region of interest (see Methods) with the cortical area that showed significant responses ($p < .05$ uncorrected) to the contrast of all visual events (F1B1, F1B2, F2B1 and F2B2) versus null events.

veridical perception of two flashes on many F2B1 trials. We found that brain activity evoked in human V1 on these fusion illusion trials (F2B1-Fusion Illusion) was significantly lower than on physically identical trials where no illusion was reported (F2B1-no Illusion). In agreement with previous findings (Bhattacharya et al., 2002; Watkins et al., 2006) we also demonstrated that activity on fission illusion trials (F1B2-Fission Illusion) was significantly higher than on physically identical trials where no illusion was seen (F1B2-no Illusion). This modulation of activity in association with illusory perception did not reflect differences in eye position or eye movements on different trials. Thus, perception of either the fission or fusion illusion caused opposite effects on activity in primary visual cortex. When two flashes were presented but one perceived, activity was decreased; but when one presented and two perceived, activity was increased. The level of cortical activity in

V1 was therefore associated with conscious visual perception rather than the physically present stimulus. We found modulation of activity by illusory multisensory perception only in the stimulus responsive area of primary visual cortex. This demonstrates that an auditory effect in primary visual cortex is specific to the area representing the visual stimulus, and reflects a modulatory influence on visual stimulation.

Critically, these divergent effects on V1 activity that follow perception cannot be explained by a general attentional effect, nor a response of early visual areas to a match between physically present stimuli and perception. For the fission multisensory illusion reported here, visual evoked potentials and fields associated with the illusory perception are modified at a short latency (Bhattacharya et al., 2002; Shams et al., 2001, 2005) consistent with generators in early visual cortex (although note that temporally early effects

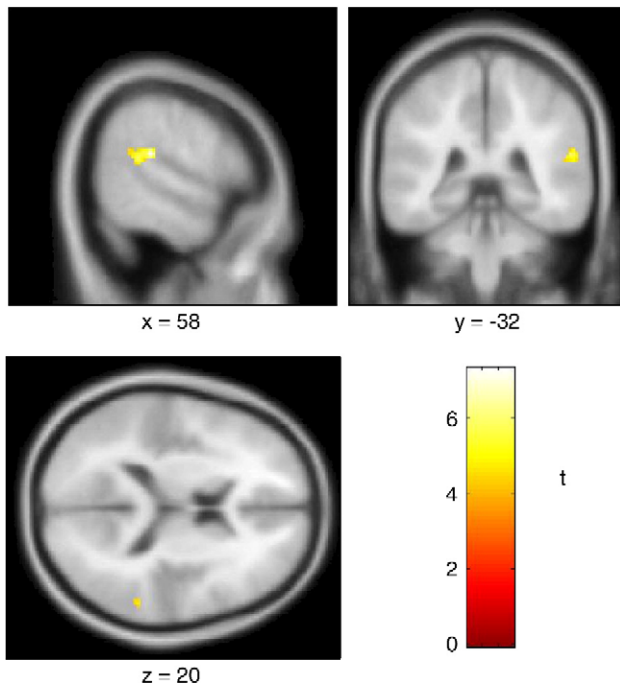


Fig. 3. Cortical areas activated by multisensory illusory perception outside retinotopic cortex. Shown in the figure are cortical loci outside retinotopic cortex where event-related activity was significantly greater during F2B1-Fusion Illusion trials compared to F2B1-no Illusion trials ($p < .05$, corrected for multiple comparisons; see also Results). Activated cortical loci in the right superior temporal sulcus projected onto a T1 template image in the stereotaxic space of Talairach and Tournoux (1988).

reported from earlier studies do not necessarily translate into anatomically early effects, such as generators in the early visual cortex). In addition, our previous functional imaging work has demonstrated increased cortical activity in V1 in association with illusory perception of an additional visual flash (Watkins et al., 2006). Our new findings extend this earlier work by demonstrating conclusively that activity in V1 follows multisensory perception.

The general finding that V1 activity can be more closely related to conscious visual experience rather than physical stimulation is recognised in unisensory studies. For example, activity evoked in human V1 by a visual stimulus briefly presented at the contrast detection threshold is higher on trials when subjects successfully detect it than when they fail to do so. Moreover, when subjects falsely perceive the presence of a low-contrast stimulus on trials when the stimulus was physically absent (false alarms), V1 activity is similar to that on trials where subjects correctly report the physical presence of a stimulus (Ress and Heeger, 2003). The present findings show that such an association of V1 activity with conscious perception extends to suprathreshold visual stimuli and to changes in visual perception brought about by multisensory stimulation.

Our data do not precisely define how the association of V1 activation with illusory visual perception occurs, nor whether the modulation of V1 activity we observed plays a causal role in the generation of the illusion. Primary visual cortex receives projections from at least 12 areas belonging to the visual cortex (Felleman and Van Essen, 1991). Recently more distant projections have been described from areas in the ventral (Distler et al., 1993) and dorsal visual pathways and from the lateral intraparietal area (Boussaoud et al., 1990; Rockland and Van Hoesen, 1994). Several

recent papers have used tracer injections to demonstrate projections from primary auditory cortex, auditory association areas and the superior temporal polysensory area (STP) to the area of primary visual cortex representing the peripheral visual field (Clavagnier et al., 2004; Falchier et al., 2002; Rockland and Van Hoesen, 1994). The function of these projections to V1 has been the subject of much debate they may serve to enhance perceptual capabilities; for example the addition of an auditory signal to a visual signal leads to improved detection compared to a visual signal alone (Bolognini et al., 2005; Frassinetti et al., 2002; Gondan et al., 2005; Miller, 1982; Molholm et al., 2002; Schroger and Widmann, 1998). Thus, it is possible that these direct connections mediate the changes in V1 activity that we observed. This is consistent with previous findings showing that multisensory influences also extend to the earliest stages of cortical processing (Ghazanfar and Schroeder, 2006). Interestingly, recent studies examining somatosensory–auditory multisensory integration in primary auditory cortex show that auditory input to A1 occurs via feedforward projections from the thalamus and it is possible that similar low level thalamic connections may extend to auditory–visual multisensory integration (Lakatos et al., 2007).

When subjects experienced the fusion illusion, activity was increased in the right superior temporal sulcus (STS) and decreased in primary visual cortex. We have thus now found evidence that the right STS is involved in both the fission and fusion illusions. The area of the right STS activated in the fission illusion is posterior to the cortical area involved in the fusion illusion. However, the size of the clusters and spatial smoothness of our data mean that it is not at present clear whether these activations reflect two distinct cortical loci. However, as regions of the right STS show a similar response for two illusions that are both perceptually very different and exhibit very different activation patterns in V1, these data suggests that the right STS may not be playing a causal role in generating the illusory perception. The low temporal resolution of fMRI signals mean that we cannot determine whether the STS activation we observed was casually related to the changes in V1 activity, or a later effect. However, we speculate that the STS response may occur later and represent a response to the matching of auditory and visual perception (i.e. the perception of F1B2 Fission illusion would be effectively F2B2 compared to the non illusion perception of F1B2). Such a speculation would be consistent with both the divergent effects of the fusion and fission illusions on activity in primary visual cortex, and previous studies demonstrating early audiovisual integration (Bhattacharya et al., 2002; Giard and Peronnet, 1999; Shams et al., 2001). The STS has been consistently associated with integration between visual and auditory stimulation (Barraclough et al., 2005; Beauchamp et al., 2004; Beauchamp, 2005; Calvert et al., 2000; Ghazanfar et al., 2005; Olson et al., 2002; Schroeder and Foxe, 2002). Interestingly, a recent study has shown that the STS is involved in multisensory associative learning (Tanabe et al., 2005). Further research will be needed to elucidate the precise role of the STS in these multisensory illusions.

Conclusion

We found that fMRI signals from stimulus-responsive regions of human primary visual cortex closely corresponded to multisensory perception for both ‘fission’ and ‘fusion’ illusions. Moreover, when auditory stimulation gave rise to an illusory change in perceptual experience this was associated with increased activity in the right superior temporal sulcus.

Acknowledgment

This work was supported by the Wellcome Trust.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2007.05.027.

References

- Andersen, T.S., Tiippana, K., Sams, M., 2004. Factors influencing audiovisual fusion and illusion. *Brain Res. Cogn. Brain Res.* 21, 301–308.
- Barracough, N.E., Xiao, D., Baker, C.I., Oram, M.W., Perrett, D.I., 2005. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J. Cogn. Neurosci.* 17, 377–391.
- Beauchamp, M.S., 2005. See me, hear me, touch me: multisensory integration in lateral occipital–temporal cortex. *Curr. Opin. Neurobiol.* 15, 145–153.
- Beauchamp, M.S., Argall, B.D., Bodurka, J., Duyn, J.H., Martin, A., 2004. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat. Neurosci.* 7, 1190–1192.
- Bhattacharya, J., Shams, L., Shimojo, S., 2002. Sound-induced illusory flash perception: role of gamma band responses. *NeuroReport* 13, 1727–1730.
- Bolognini, N., Frassinetti, F., Serino, A., Ladavas, E., 2005. “Acoustical vision” of below threshold stimuli: interaction among spatially converging audiovisual inputs. *Exp. Brain Res.* 160, 273–282.
- Boussaoud, D., Ungerleider, L.G., Desimone, R., 1990. Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J. Comp. Neurol.* 296, 462–495.
- Calvert, G.A., Campbell, R., Brammer, M.J., 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657.
- Clavagnier, S., Falchier, A., Kennedy, H., 2004. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn. Affect. Behav. Neurosci.* 4, 117–126.
- Distler, C., Boussaoud, D., Desimone, R., Ungerleider, L.G., 1993. Cortical connections of inferior temporal area TEO in macaque monkeys. *J. Comp. Neurol.* 334, 125–150.
- Falchier, A., Clavagnier, S., Barone, P., Kennedy, H., 2002. Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759.
- Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Foxe, J.J., Schroeder, C.E., 2005. The case for feedforward multisensory convergence during early cortical processing. *NeuroReport* 16, 419–423.
- Frassinetti, F., Pavani, F., Ladavas, E., 2002. Acoustical vision of neglected stimuli: interaction among spatially converging audiovisual inputs in neglect patients. *J. Cogn. Neurosci.* 14, 62–69.
- Ghazanfar, A.A., Schroeder, C.E., 2006. Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285.
- Ghazanfar, A.A., Maier, J.X., Hoffman, K.L., Logothetis, N.K., 2005. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.* 25, 5004–5012.
- Giard, M.H., Peronnet, F., 1999. Auditory–visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490.
- Gondan, M., Niederhaus, B., Rosler, F., Roder, B., 2005. Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. *Percept. Psychophys.* 67, 713–726.
- Josephs, O., Deichmann, R., Turner, R., 2000. Trajectory measurement and generalised reconstruction in rectilinear EPI. *Proc. ISMRM* 8, 1517.
- Lakatos, P., Chen, C.M., O’Connell, M.N., Mills, A., Schroeder, C.E., 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292.
- McCurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. *Nature* 264, 746–748.
- Miller, J., 1982. Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* 14, 247–279.
- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2002. Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res. Cogn. Brain Res.* 14, 115–128.
- Mottonen, R., Krause, C.M., Tiippana, K., Sams, M., 2002. Processing of changes in visual speech in the human auditory cortex. *Brain Res. Cogn. Brain Res.* 13, 417–425.
- Murray, M.M., Michel, C.M., Grave, d.P., Ortigue, S., Brunet, D., Gonzalez, A.S., Schnider, A., 2004. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage* 21, 125–135.
- Murray, M.M., Molholm, S., Michel, C.M., Heslenfeld, D.J., Ritter, W., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2005. Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex* 15, 963–974.
- Olson, I.R., Gatenby, J.C., Gore, J.C., 2002. A comparison of bound and unbound audio–visual information processing in the human cerebral cortex. *Brain Res. Cogn. Brain Res.* 14, 129–138.
- Ress, D., Heeger, D.J., 2003. Neuronal correlates of perception in early visual cortex. *Nat. Neurosci.* 6, 414–420.
- Rockland, K.S., Van Hoesen, G.W., 1994. Direct temporal–occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cereb. Cortex* 4, 300–313.
- Schroeder, C.E., Foxe, J.J., 2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.* 14, 187–198.
- Schroger, E., Widmann, A., 1998. Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology* 35, 755–759.
- Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889–893.
- Shams, L., Kamitani, Y., Shimojo, S., 2000. Illusions. What you see is what you hear. *Nature* 408, 788.
- Shams, L., Kamitani, Y., Thompson, S., Shimojo, S., 2001. Sound alters visual evoked potentials in humans. *NeuroReport* 12, 3849–3852.
- Shams, L., Kamitani, Y., Shimojo, S., 2002. Visual illusion induced by sound. *Brain Res. Cogn. Brain Res.* 14, 147–152.
- Shams, L., Ma, W.J., Beierholm, U., 2005. Sound-induced flash illusion as an optimal percept. *NeuroReport* 16, 1923–1927.
- Stein, B.E., London, N., Wilkinson, L., Price, D., 1996. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J. Cogn. Neurosci.* 8, 497–506.
- Talairach, J., Tournoux, P., 1988. *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme, Stuttgart, Germany.
- Tanabe, H.C., Honda, M., Sadato, N., 2005. Functionally segregated neural substrates for arbitrary audiovisual paired-association learning. *J. Neurosci.* 25, 6409–6418.
- Teo, P.C., Sapiro, G., Wandell, B.A., 1997. Creating connected representations of cortical gray matter for functional MRI visualization. *IEEE Trans. Med. Imag.* 16, 852–863.
- Wandell, B.A., Chial, S., Backus, B.T., 2000. Visualization and measurement of the cortical surface. *J. Cogn. Neurosci.* 12, 739–752.
- Watkins, S., Shams, L., Tanaka, S., Haynes, J.D., Rees, G., 2006. Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage* 31, 1247–1256.
- Weiskopf, N., Sitaram, R., Josephs, O., Veit, R., Scharnowski, F., Goebel, R., Birbaumer, N., Deichmann, R., Mathiak, K., in press. Real-time functional magnetic resonance imaging: methods and applications. *Magnetic Resonance Imaging*.