Perception of the touch-induced visual double-flash illusion correlates with changes of rhythmic neuronal activity in human visual and somatosensory areas

Joachim Lange\textsuperscript{a, b}, Robert Oostenveld\textsuperscript{a, c}, Pascal Fries\textsuperscript{a, c}

\textsuperscript{a} Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands
\textsuperscript{b} Medical Faculty, Institute for Clinical Neuroscience and Medical Psychology, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany
\textsuperscript{c} Ernst Strüngmann Institute (ESI) in Cooperation with Max Planck Society, Deutschordenstr. 46, 60528 Frankfurt, Germany

Corresponding author:
Joachim Lange
Medical Faculty, Institute for Clinical Neuroscience and Medical Psychology, Universitätsstraße 1, 40225 Düsseldorf, Germany
Tel: +492118113074
Fax: +492118113015
e-mail: joachim.lange@med.uni-duesseldorf.de

Abstract

A single brief visual stimulus accompanied by two brief tactile stimuli is frequently perceived incorrectly as two flashes, a phenomenon called double-flash illusion (DFI). We investigated whether the DFI is accompanied by changes in rhythmic neuronal activity, using magnetoencephalography in human subjects. Twenty-two subjects received visuo-tactile stimulation and reported the number of perceived visual stimuli. We sorted trials with identical physical stimulation according to the reported subjective percept and assessed differences in spectral power in somatosensory and occipital sensors. In DFI trials, occipital sensors displayed a contralateral enhancement of gamma-band (80–140 Hz) activity in response to stimulation. In somatosensory sensors, the DFI was associated with an increase of spectral power for low frequencies (5–17.5 Hz) around stimulation and a decrease of spectral power in the 22.5–30 Hz range between 450 and 750 ms post-stimulation. In summary, several components of rhythmic activity predicted variable subjective experience for constant physical stimulation. Notably, the enhanced occipital gamma-band activity during DFI was similar in time and frequency extent to the somatosensory gamma-band response to tactile stimulation. We speculate that the DFI might therefore occur when the somatosensory gamma-response is transmitted to visual cortex. This transmission might be supported by the observed modulations in low-frequency activity.
Research highlights
- Presentation of one visual stimulus accompanied by two tactile stimuli.
- Frequent illusionary perception of a second visual stimulus (double-flash illusion).
- Changes of oscillatory activity in different frequency bands predict illusion.
- Spontaneous fluctuations of attention might influence perception of illusion.
- Effective visuo-tactile transmission of oscillatory activity might underlie illusion.

Keywords
Illusion; Oscillation; MEG; Somatosensory; Visual
Introduction
Rhythmic neuronal activity has been proposed as a crucial factor for communication among neuronal groups (Canolty et al., 2006, Fries, 2005, Lachaux et al., 2005, Salinas & Sejnowski, 2001, Tallon-Baudry et al., 2004, Tallon-Baudry et al., 2005 and Womelsdorf et al., 2007). Putative specific roles of coherent rhythmic activity include top-down modulation (Engel et al., 2001) or the effective long-range communication of neuronal groups (Gross et al., 2004, Lachaux et al., 2005 and Varela et al., 2001). Synchronization of rhythmic activity has been found in cortico-cortical (Gross et al., 2004) and cortico-spinal networks (Schoffelen et al., 2005).

Several studies argue that long-range communication through rhythmic neuronal activity might provide a mechanism also for cross-modal interaction. Multisensory stimuli evoke stronger rhythmic activity in the gamma-band than unisensory stimuli (Sakowitz et al., 2001 and Senkowski et al., 2007). A recent intracranial study in monkeys revealed that somatosensory input in auditory cortex resets the phase of ongoing auditory cortical oscillations which leads to enhanced processing capabilities of audio-tactile stimuli (Lakatos et al., 2007). Evidence for a role of rhythmic neuronal activity for cross-sensory interaction has also been found in a cross-modal illusion (Bhattacharya et al., 2002 and Mishra et al., 2007). The authors presented a briefly flashed visual stimulus accompanied by two brief auditory stimuli. This stimulation triggers, in a subset of trials, the illusory perception of a second visual stimulus, i.e. the double-flash illusion (DFI) (Shams et al., 2000). Recently, it has been demonstrated that the perception of a visual stimulus can also be altered by tactile stimulation in a similar way (Violentyev et al., 2005). Both, the auditory induced and the tactile induced DFI, do not merely shift unimodal perception quantitatively (e.g. to a higher perceived luminance), but rather alter the phenomenological quality of the percept. Such illusions in which identical stimulation can lead to categorically different visual perception are rare but offer an intriguing opportunity to study cross-modal interactions and perception in general (Haynes et al., 2005, Leopold et al., 2002 and Wilke et al., 2006).

The neurophysiological basis of the DFI has been investigated for the audio-visual DFI. These EEG studies found among other results that gamma band activity in occipital areas was enhanced for trials in which subjects perceived the illusion (Bhattacharya et al., 2002 and Mishra et al., 2007). The neurophysiological basis of the touch-induced visual DFI has so far not been investigated. Since rhythmic neuronal activity has been proposed as a general mechanism for cross-modal integration, the objective of the present study was to investigate its putative role for the dynamic interaction of somatosensory and visual areas in the visuo-tactile DFI.

Methods
Subjects
Twenty-two right-handed volunteers (12 female, mean age (± SD) 22.3 ± 2.7 years) without any known history of neurological disorders participated in this study. All participants had normal or corrected-to-normal vision. The experiment was approved by the local ethics committee, and each subject gave written informed consent prior to the experiment, according to the Declaration of Helsinki.

Stimuli, tasks, and procedure
Subjects were lying comfortably in supine position with their head placed in the horizontally positioned MEG helmet. Visual stimuli were presented using an LCD projector located outside the magnetically shielded room and back-projected onto a translucent screen via a mirror-system. The vertical refresh rate of the LCD projector was 60 Hz. The screen was positioned 70 cm in front of the subjects so that they could comfortably view the visual stimuli.

The experimental paradigm and the stimuli are illustrated in Fig. 1. Each trial started with the presentation of a fixation point (Gaussian of diameter 0.5°, luminance 7 cd/m²) in the middle of the screen. Subjects maintained fixation on this dot throughout the entire trial. After 700 ms, a stimulation period of 90 ms started in which subjects received visual and/or tactile stimulation. Visual stimulation consisted of uniform gray discs (2.5° diameter, average luminance over subjects 2.3 cd/m², see below for details about the luminance of the visual stimuli) presented for 16 ms, 17° left of the fixation dot. Visual stimuli were presented eccentrically because this increased the likelihood to experience the visual illusion (Shams et al., 2000).

![Figure 1 Schematic illustration of the experimental paradigm and setup. A, schematic overview of the setup. Subjects fixated a central dot while they received tactile stimulation on their left index finger and/or visual stimulation. Visual stimuli consisted of a gray disc presented at 17° eccentricity in the left visual hemifield, on the horizontal meridian. Subjects responded by pressing one of three buttons with their right hand. B, illustration of a single trial and of the different experimental conditions. All nine conditions were presented in random order within a block. The entire experiment consisted of 100 blocks.](image)

Tactile stimulation was given through a piezoelectrical Braille stimulator (Metec, Stuttgart, Germany) (Bauer et al., 2006) that was taped to the subjects’ left index finger. The Braille stimulator consisted of a 2 × 4 matrix of pins and was driven using custom-built electronic circuitry. For the tactile stimulation, all eight pins were lifted simultaneously by 2 mm, stayed elevated for 30 ms, and were then lowered again.

The number of visual and/or tactile stimuli could be zero, one, or two. All nine possible visuo-tactile combinations were used (Fig. 1B). We will address the different conditions as “vxty” for a condition with x visual and y tactile stimuli. Thus, stimuli could be either single unimodal (v1t0, v0t1), double unimodal (v2t0, v0t2) with onsets of both stimuli separated by 60 ms, or bimodal combinations of
tactile and visual stimuli. Bimodal visuo-tactile stimuli could either appear synchronously (v1t1, v2t2). In the latter condition (v2t2) the onsets of both pairs of visuo-tactile stimuli were separated by 60 ms. Or one tactile stimulus was presented between two visual stimuli (v2t1) or vice versa (v1t2). Finally, one condition did not contain any stimulation at all (v0t0).

The stimulation period was followed by a post-stimulus period of 1000 ms in which only the fixation dot was visible. Then, the contrast of the fixation dot was increased by 40% indicating the response period. Subjects were asked to report the number of perceived visual stimuli and to ignore the tactile stimulations. Responses were given by pressing one of three buttons with fingers of the right hand (Fig. 1A) on an optical button box (Lumitouch, Photon Control Inc., Burnaby, Canada). Button presses were counterbalanced for the critical responses (i.e. one or two stimuli). For half of the subjects the button configuration was: thumb press: zero stimuli; index finger: one stimulus; middle finger: two stimuli; for the other half of the subjects the button configuration was: thumb press: zero stimuli; middle finger: one stimulus; index finger: two stimuli. Additionally, subjects reported their confidence by pressing the respective button once (confident) or twice (not confident). The fraction of not-confident trials was below 5%. Thus, confident and not-confident trials were collapsed for the analysis. Since confidence ratings were high for each subject, we conclude that the results express subjects' subjective perception rather than indecisiveness in a two-alternative choice task, in agreement with a recent direct investigation of this issue (McCormick and Mamassian, 2008).

After the subject's response, or maximally 3000 ms after the increase in fixation dot luminance, the next trial started. Only responses given during the response period were analyzed. Trials in which the response was given before the response period were discarded from analysis.

Both, the prestimulus and the post-stimulus intervals were constant for the following reasons:

1) The study focuses on spectral power. Since power is a rectified quantity, anticipatory perturbations would be visible in the average with both, a constant and a varying stimulus/response timing.

2) By varying the time of stimulus onset, we would have created a time-varying hazard rate and corresponding expectation and thereby had created more variance of neuronal activity at stimulus onset.

3) The main focus of our study was on illusion effects, i.e. on perceptual contrasts between conditions of identical physical stimulation. A fixed time interval minimizes variance due to anticipation.

For the analysis of the effects of sensory stimulation per se, a post-stimulus period is compared to a pre-stimulus baseline. This comparison is not immune to anticipatory signals. However, the stimulation contrasts are not central to the study. They are only used to constrain the sensors (regions of interest) for the subsequent investigation of the illusion contrasts. Furthermore, the stimulation contrasts basically confirm existing findings and are thus very unlikely caused purely by expectation.

Previous studies reported a high inter-subject variability of illusory trials despite identical stimulations for all subjects (Mishra et al., 2007, Shams et al., 2000 and Violentyev et al., 2005). To ensure a balanced ratio of illusory and non-illusory trials within each subject, we adapted the
luminance of the visual stimuli for each subject individually prior to the experiment by a staircase method. These pre-tests also served as training trials for the subjects and lasted about 5 min. The resulting luminance of the visual stimuli was $2.3 \pm 0.2 \text{ cd/m}^2$ (mean $\pm$ SEM).

To increase statistical power for the investigation of the illusory effect, the condition v1t2 was presented twice as often as the other conditions. One pseudorandom sequence of all conditions (including two times v1t2) constituted one block. After ten blocks, subjects were allowed to take a short break. Overall, the experiment consisted of 100 blocks, resulting in 100 trials for each condition (200 trials for condition v1t2) and a recording session of ~ 1 h. Subjects were instructed to blink only during the response period or during the breaks and to press the buttons only during the response periods.

The sound generated by the Braille cells was strongly attenuated by encapsulating them into foam. Additionally, a very weak residual sound was masked by presenting subjects with auditory white noise via pneumatic earphones. Behavioral tests preceding the experiment confirmed that subjects did not hear the clicking of the Braille cells.

Stimuli were controlled using the software “Presentation” (Neurobehavioral Systems, Albany).

Data acquisition
MEG recordings

Electromagnetic brain activity was recorded using a whole head 151-channel or 275-channel MEG system (CTF systems Inc., Port Coquitlam, Canada). Data from the 275-channel system were interpolated to a common 151-channel template using a procedure that was also used to compensate for differences in subjects' head position (see paragraph 2.4.1 on Preprocessing for details). The system was moved to a horizontal position and subjects were recorded in supine position. We recorded vertical and horizontal eye movements simultaneously, by measuring the electroocculogram (EOG) through electrodes placed below and above the left eye and on the outer sides of each eye. MEG/EOG data were low-pass filtered at 300 Hz and sampled continuously at a rate of 1200 Hz. Subjects' head position relative to the gradiometer array was determined before and after the recording session by measuring the position of reference coils placed at the subjects' nasion and at the left and right ear canals.

Structural MRI
For each subject, a full-brain high-resolution anatomical MR image was acquired on a 1.5T or 3T whole-body scanner (Siemens, Erlangen, Germany) using a volume head coil for radio frequency transmission and signal reception. We applied standard T1-weighted sequences. MRI reference markers were placed at the subjects' nasion and at the left and right ear canals for alignment of the MEG and MRI coordinate systems.

Data analysis
Preprocessing
Data were analyzed using the FieldTrip software (http://fieldtrip.fcdonders.nl), an open source Matlab toolbox for neurophysiological data analysis developed at the Donders Institute for Brain, Cognition and Behaviour, Nijmegen.

Power line noise was removed from the continuous data using a Fourier transformation of 10-s long signal periods and subtracting the 50, 100 and 150 Hz components. This was done separately for all 10-s periods around all periods of interest.

Using a semi-automatic routine, segments contaminated by eye movements or blinks, artifacts caused by muscle activity, and jump artifacts in the MEG signal caused by the SQUID electronics were removed. If the length of a trial was smaller than 800 ms, the entire trial was removed. All trials in which subjects responded too early (i.e. during the 1000 ms post-stimulus period after stimulus presentation, see Fig. 1), or trials in which no response was given, were discarded.

Differences in subjects' head positions with respect to the MEG sensors may cause smearing of activity when constructing the grand-average topography over subjects. To compensate for this and to interpolate the data from the 275-channel and the 151-channel system, the individual subjects' MEG data were interpolated to a common 151-channel-template position for the MEG sensors with respect to the head. Each subjects' head shape was reconstructed from the MRI and a superficial layer of dipoles was placed 1 cm beneath the inner skull surface, approximately in the sulcal gray matter. This dipole layer consisted of 642 evenly spaced nodes covering the whole surface, with a regional source (i.e. free orientation) at every node. The leadfield matrix for all combined sources was computed using an individual forward model for every subject (Nolte, 2003). The strength of each of the 3 × 642 dipole components was estimated using a minimum norm estimate (MNE). The MNE was done using the More-Penrose pseudoinverse of the leadfield matrix by computing the singular value decomposition and regularized by truncating the singular values of the leadfield matrix at 1/1000 of the maximal singular value. The estimated strength of these dipoles in the gray matter was subsequently used to compute the field distribution at the location of each sensor in the template gradiometer array, again using the individual forward model. The placement of the dipole layer in grey matter ensures that the forward computed field distribution resembles the actual distribution of the field of the true underlying sources. This method is robust and yields accurate results, regardless of the position of the field generator (Knösche, 2002). Subsequently, planar gradients for a given sensor were estimated from the axial field distribution and computed separately in vertical and horizontal direction by comparing the field at that sensor with its neighbors (Bastiaansen and Knösche, 2000). An advantage of the planar gradient transformation over axial gradient data is that the signal amplitude is largest directly above a source. This is particularly advantageous when interpreting distributions of spectral power at the sensor level.

**Spectral analysis**

We analyzed the data in the frequency domain. Power spectra were computed by using a short time discrete Fourier transform on temporal windows sliding in 25-ms steps. Two frequency ranges were analyzed separately: A low-frequency band (5–40 Hz) was analyzed with analysis windows of 400 ms length, tapered with a single Hanning window resulting in a spectral smoothing of roughly ± 2.5 Hz. A high-frequency band (40–150 Hz) was analyzed with analysis windows of 200 ms length. For
the high-frequency band, we applied a multi-taper approach to the respective analysis windows to optimize spectral concentration over the frequency of interest (Mitra and Pesaran, 1999). We applied eleven Slepian tapers resulting in a spectral smoothing of ± 30 Hz. Spectral power was first estimated per trial and taper and then averaged across trials and tapers.

The long window for the low-frequency band was chosen, because known, physiological low-frequency bands are relatively narrow and closely spaced. We therefore aimed at a spectral resolution in the low-frequency range of roughly ± 2.5 Hz. A spectral resolution of 2.5 Hz requires analysis windows of at least the inverse of 2.5 Hz, i.e. 0.4 s.

In the higher frequency range, frequency bands of physiological rhythms are broader and spaced more far apart. We could therefore afford to use a window that was half the size used for the low frequencies, i.e. 0.2 s. A further reduction in window length would have been problematic. (1) It would have rendered the windows of temporal support for low and high frequencies even more disparate. (2) It would have created problems in the group analysis when different subjects had slightly different times of spectral perturbations.

We defined regions of interest (ROIs) in sensor space, in which we analyzed the spectral power. The choice of ROIs was motivated by our focus on sensory representations. We therefore defined the ROIs as the region showing the most consistent spectral perturbations of rhythmic activity in response to tactile/visual stimulation. First, all trials with purely tactile stimulation (v0t1 and v0t2) were pooled. These trials showed clear spectral components in the low-, mid- and high-frequency bands (Fig. 3A). For each time–frequency band, we determined the topographical distribution (Fig. 3B). We normalized and rectified the amplitudes of the three resulting patterns (i.e. response amplitudes were normalized to values between 0 and 1), averaged these patterns, and determined the 10 sensors which captured the maxima of the averaged pattern. The spectral patterns overlapped in 10 MEG sensors in the right central region, contralateral to tactile stimulation (RC13, RC14, RC15, RC21, RC22, RC23, RC24, RC31, RC32, RP34). This set of sensors defined the somatosensory region of interest.

Secondly, all trials with purely visual stimulation (v1t0 and v2t0) were pooled. These trials showed four clear spectral components in the low-, mid- and high-frequency bands (Fig. 4A). In line with the analysis of the somatosensory ROI, we averaged the normalized and rectified the four resulting topographical representations and determined the 10 sensors which captured the maxima of the averaged pattern (Fig. 4B). The spectral patterns overlapped in 10 MEG sensors in the right and central occipital region, contralateral to visual stimulation (RO21, RO22, RO31, RO32, RO33, RO41, RO42, ZO01, LO31, LO32). This set of sensors defined the occipital region of interest.

**Statistical analysis**

We tested for effects of sensory stimulation and for effects of subjective perception.

To test for the overall effect of tactile and visual stimulation, all trials with purely tactile stimulation (v0t1 and v0t2) were pooled into a new tactile-only condition, and all trials with purely visual
stimulation (v1t0 and v2t0) were pooled into a new visual-only condition. The response to sensory stimulation was assessed by comparing, within those unimodal stimulation conditions, the post-stimulus period to the pre-stimulus baseline period (−300 ms to −100 ms before stimulus onset, collapsed across time). To test for an effect of the DFI, we sorted the v1t2 trials into trials with DFI and without DFI.

For each subject separately, we averaged the spectral power over the sensors of interest (see Spectral analysis). We visually inspected the power spectra and defined broad time–frequency regions of interest that captured visible perturbations, but restricted the time–frequency range over which multiple comparison correction was performed later. Those broad time–frequency regions correspond to the time–frequency ranges shown in the figures. The two conditions were compared with an independent samples t-test (i.e. “post-stimulus versus baseline,” or “DFI versus non-DFI”). This was done for each subject and each time–frequency pixel and therefore resulted in time–frequency t-maps. The subsequent group level statistic determined whether inside those regions of interest, there were time–frequency clusters with effects that were significant at the random effects level after correcting for multiple comparisons along both the time and the frequency dimension.

The group level statistic used those t-maps as inputs and determined their consistency across subjects. The null hypothesis was that the data from the two conditions were not different and could therefore be exchanged. We therefore tested exchangeability using a non-parametric permutation approach (Maris and Oostenveld, 2007). We choose this approach for several reasons: First, it is free of assumptions about the underlying distributions. Second, it is not affected by the fact that there was partial dependence (due to overlap) between neighboring time–frequency pixels. Third, it offers an elegant way to correct for multiple comparisons. The procedure was as follows:

1) We defined the test-statistic to be the t-value pooled across subjects, i.e. the sum of the individual subjects’ t-values divided by the square root of the number of subjects. This test-statistic was determined per time–frequency pixel.
2) The pooled t-map was thresholded with a non-multiple comparison corrected a priori t-value threshold of 1.96, corresponding to a parametric two-sided paired t-test with 5% false positives.
3) This resulted in clusters of adjacent time–frequency pixels for which we determined the sum of the test statistic. This sum was defined to be our cluster-level test statistic.
4) We performed 5000 randomizations. In each randomization, we selected a random subset of subjects. For those subjects, the t-values were inverted, i.e. all time–frequency t-values were multiplied with minus one. This is equivalent to an exchange of the two conditions and therefore implements a new dataset under the null-hypothesis of exchangeability (Maris and Oostenveld, 2007).
5) For each of the 5000 randomization, steps (1) through (3) were repeated.
6) For each randomization, only the maximal and the minimal cluster-level test statistic across all clusters were retained and placed into two histograms, which we address as maximum (or minimum, respectively) cluster-level test statistic histograms.
7) After all 5000 randomizations, we determined, for each cluster from the observed data, the fraction of the maximum (minimum) cluster-level test statistic histogram that was greater (smaller) than the cluster-level test statistic from the observed cluster. The smaller of the two
fractions was retained and divided by 5000, giving the multiple comparisons corrected significance thresholds for a two-sided test.

Please note that this procedure implements a random effect analysis, because the randomization (condition exchange, i.e. \( t \)-value inversion) was done at the level of subjects. Had we simply thresholded the pooled \( t \)-values, this had implemented a fixed effect analysis, but rather we used those pooled \( t \)-values only as a well normalized input to the group level random effect analysis.

To analyze the contribution of time-locked (evoked) components, we sorted the \( v1t2 \) trials into DFI and non-DFI trials and averaged the trials in the time-domain. To optimize the analysis of time-locked data, averaging was performed before computation of planar fields. Additionally, we computed the spectral power of these time-locked data. Statistical testing was performed separately in the time and frequency domain.

**Correlation analysis**

The main, early post-stimulus DFI effects were found to be enhancements of somatosensory low-frequency and visual gamma-band activity. For those effects, we tested whether the somatosensory and visual effects were correlated across trials. For each of the effects, those time–frequency points were selected that showed significant DFI effects on the group level. Subsequently, for each subject and trial, the power in those time–frequency regions was expressed as percent change relative to baseline. Within each subject, we then determined the Spearman rank correlation between visual gamma and somatosensory low-frequency across trials. The resulting correlation coefficients were tested against a null hypothesis of no correlation, using a one-sample \( t \)-test across subjects.

**Eye movements**

Although trials contaminated by eye movements have been discarded during preprocessing, we tested for residual eye movements. To analyze directly whether eye-movements differed between DFI trials and non-illusory trials, we applied the procedure described in the section Statistical analysis for MEG-channels to the EOG-channels in the time domain. To this end, we first computed eye movements in horizontal and vertical direction and then combined them to assess overall eye movements. For the combination, we used an Euclidian measure (\( \text{EOG}_{\text{comb}} = \sqrt{\text{EOG}_{\text{vert}}^2 + \text{EOG}_{\text{hor}}^2} \)), with \( \text{EOG}_{\text{comb}} \) denoting combined EOG channels and \( \text{EOG}_{\text{vert}} \) and \( \text{EOG}_{\text{hor}} \) the vertical and horizontal channels, respectively. Statistical analysis was performed separately for horizontal, vertical and overall eye movements.

**Results**

**Behavioral results**

Subjects made negligible errors when judging the number of presented visual stimuli in six out of nine conditions (Fig. 2). These were the conditions without visual stimulation, the conditions in which the number of visual and tactile stimuli was the same and the condition in which only one visual stimulus was presented. However, trials with one visual and two tactile stimuli (\( v1t2 \) condition), were perceived as two visual stimuli in \( 45.8 \pm 0.4\% \) (mean ± SEM). Statistical analysis revealed that the presence of a second tactile stimulus in condition \( v1t2 \) significantly increased the proportion of incorrect responses (i.e. perception of illusory second flashes) compared to condition
v1t1 (t(21) = 11.9, p < .0001), i.e. confirming previous results that subjects experienced the double flash illusion (DFI) (Mishra et al., 2007, Shams et al., 2000 and Violentyev et al., 2005). To confirm that the DFI is caused by perceptual processes, we performed an additional analysis based on signal detection theory. For this analysis, we denoted correctly perceived double flashes as “hits”, correctly perceived single flashes as “correct rejections.” Trials with a single flash that were erroneously perceived as two flashes were denoted as “false alarms,” and trials with two flashes that were erroneously reported as a single flash as “misses” (McCormick & Mamassian, 2008, Violentyev et al., 2005 and Watkins et al., 2006). The presence of two tactile stimuli significantly decreased sensitivity (d’ = 1.78 ± 0.16; mean ± SEM) by 19.8% compared to the presence of one tactile stimulus (d’ = 2.19 ± 0.20, t(21) = 2.24, p = 0.036). This confirms previous results showing that the DFI is a perceptual effect and not caused by a simple shift of criterion bias (McCormick & Mamassian, 2008, Violentyev et al., 2005 and Watkins et al., 2006).

**Stimulation contrast**

To study the effect of tactile stimulation per se, all trials with purely tactile stimulation (v0t1 and v0t2) were pooled. We focused on the effects of stimulation, thus, we averaged over all (artifact free) trials irrespective of perceptual report, and we contrasted to the pre-stimulation baseline. The same procedure was applied to the purely visual stimuli (v1t0 and v2t0). Fig. 3A shows the time–frequency analysis for tactile stimulation averaged over the somatosensory sensors (see (Spectral analysis) for details). We found that tactile stimulation induced a highly significant reduction (p < 0.001) of rhythmic activity in a time–frequency cluster that had three major lobes: One between 200 and 550 ms after stimulation and in a frequency band of 15–40 Hz, a second between 300 and 700
ms and in a frequency band of 7.5–15 Hz, and a third, although less prominent lobe between 800 and 950 ms and in a frequency band of 25–40 Hz (Fig. 3A).

Figure 3 The effect of tactile stimulation of the left index finger. A, TFRs for the sensors indicated in B (black dots). t-Values were calculated separately for low and high frequencies by pooling all trials of purely tactile conditions (v0t1 and v0t2). Positive t-values indicate greater power after stimulation as compared to the pre-stimulus baseline. The negative cluster in the high frequency analysis corresponds to the negative cluster in the frequency analysis. B, separate topographies for the three clusters as highlighted in A (see Spectral analysis for details on choice of time-frequency clusters in (A) and sensors in (B)).

Significant increases of rhythmic activity were found in two time–frequency clusters: one cluster (p < 0.05) ranging from 75 to 500 ms and from 5 to 12.5 Hz and a second cluster (p < 0.05), from 100 to 300 ms and ranging from 80 to 120 Hz, i.e. in the gamma-frequency band. The upper panel of Fig. 3A also shows a strong reduction in rhythmic activity from 200 to 600 ms and for the frequencies from 40 to 70 Hz. This cluster corresponds most likely to the negative cluster observed for the low frequencies (7.5–40 Hz) and becomes visible in the high-frequency analysis because of the spectral smoothing. Therefore, we will not further discuss this effect.

We averaged the t-values of all time–frequency pixels inside significant clusters to compile topographic plots. The topographies of the significant clusters reveal that the effects peak over somatosensory areas contralateral to stimulus presentation. The stimulation effect is most spatially focused for the gamma-band. The remaining two clusters reveal in addition weak ipsilateral responses.
Fig. 4 shows the effects of visual stimulation. There was a significant enhancement of rhythmic activity (p < 0.05) with two prominent lobes: one between 200 and 400 ms and from 5 to 15 Hz, and a second lobe between 400 and 700 ms in the frequency band 12.5–40 Hz. There is a significant decrease of rhythmic activity (p < 0.05) between 400 and 900 ms and from 65 to 150 Hz. There is also a significant increase of activity between 400 and 600 ms in the frequency band 40–60 Hz. Similar to the tactile condition, this cluster corresponds most likely to the cluster observed for the low frequencies and becomes visible in the high-frequency analysis because of the spectral smoothing. Therefore, we will not further discuss this effect.

The topographies of the significant clusters reveal that the increase in the low frequencies is spatially restricted to the occipital sensors and peaks over the contralateral side. The decrease of activity in the high frequencies has a widespread topography.

**The double flash illusion (DFI) effect**

Next, we investigated the role of rhythmic activity for the perception of a second (illusory) flash. All trials with one visual and two tactile stimuli (condition v1t2) were sorted according to the perceptual report of the subjects. We compared the spectral power in the somatosensory and occipital regions for trials in which subjects reported to perceive two flashes (on average 76 trials per subject) vs. trials in which one flash was reported (on average 88 trials). Power estimates were used for further analyses directly, without subtracting a baseline. We will address results of this comparison as DFI...
effects. If not mentioned otherwise, statistical tests were performed over the entire time–frequency range. For sensors overlying somatosensory areas, we found significant DFI effects in two time–frequency clusters (Fig. 5A): One cluster (p < 0.05) of increased rhythmic activity extending from −50 to 400 ms and between 5 and 17.5 Hz, and a second cluster (p < 0.05) of reduced rhythmic activity between 450–750 ms and 22.5–30 Hz. Topographical plots of both effects revealed clear spatial foci over contralateral somatosensory sensors (Fig. 5B). We tested whether these effects were attributable to stimulus-locked components. To this end, we first averaged the signals in the time domain and then performed the spectral and statistical analysis in the time- and the spectral domain as before. This did not reveal any significant DFI effects for the stimulus-locked components (Fig. S3 and Fig. S4, see Methods (section 2.4.3.)). No significant DFI effect was observed in somatosensory sensors for higher frequencies.

Figure 5 The double flash illusion effect in somatosensory sensors. The stimulation condition was v1t2. Positive t-values indicate greater power in trials in which subjects perceived an illusory second flash as compared to trials in which they perceived the veridical single flash. Time–frequency regions that are revealed by the semitransparent white mask were significant after multiple comparison correction. Otherwise, the format is as in Fig. 3. A, TFRs for sensors indicated in B. B, separate topographies for both significant clusters as highlighted in A. The colorbar applies to both topographies.

By contrast, in sensors over visual cortex, the main DFI effect was an increase of rhythmic activity between 75 and 300 ms and in the frequency range 80–140 Hz, i.e. in the high gamma-band (Fig. 6A). Visual inspection of the power spectra revealed that high-frequencies were dominated by a strong decrease in the range 40–70 Hz and that high gamma-band effects were only visible in the post-stimulus period (Fig. S1). Therefore, the statistical analysis of the high-frequency effects in the visual cortex was restricted to 0–900 ms after stimulation and to the frequency band 70–150 Hz. The topography of this effect revealed a focus over contralateral visual cortex, with additional peaks over parietal and left frontal areas (Fig. 6B). As in the analysis over somatosensory areas, there was no
significant DFI effect for the stimulus-locked components (Fig. S3 and Fig. S5, see Statistical analysis).

![Figure 6](image.png)

Figure 6 The double flash illusion effect in visual sensors. Same analysis and format as Fig. 5. A, TFRs for sensors indicated in B. B, separate topography for significant cluster as highlighted in A.

Visual inspection of the increase of gamma-band activity for DFI-trials (Fig. 6) and the stimulus-driven gamma-band activity (Fig. 4) revealed that the DFI-effect occurred slightly earlier in time, in a slightly higher frequency-band, and also topographically more central than the stimulus-driven gamma effect in response to unimodal visual stimulation.

In conditions v2t0 and v2t1 subjects frequently missed one visual stimulus, i.e. they reported seeing only one flash (Fig. 2). Similar to the analyses of the DFI effect, we sorted the trials in conditions v2t0 and v2t1 sorted according to the perceptual report of the subjects and compared the spectral power in the somatosensory and occipital regions for trials in which subjects reported to perceive two flashes vs. trials in which one flash was reported. We did not find any significant differences between both perceptual reports.

The main early DFI effects were enhancements of somatosensory low-frequency and visual gamma-band activity. To test for a potential relation between those processes, we calculated the correlation across trial-by-trial variations in those effects (see Methods for details). We found the power enhancements in somatosensory low-frequency and visual gamma-band activity to be positively correlated (positive Spearman rank correlation in 21 of 22 subjects (t-test across subjects: t(21) = 8.4, p < 0.001).
Additional analyses compared DFI and non-DFI trials in the EOG-signal but did not show any significant differences in either horizontal, vertical or overall eye-movements (Fig. S2; see Eye movement for details).

Discussion
We studied rhythmic neuronal activity in humans during visual and/or tactile stimulation and in relation to the induced visual percept. Isolated tactile stimulation induced enhanced low-frequency and high-gamma frequency activity, as well as reduced alpha- and beta-band activity over somatosensory cortex contralateral to stimulation and very similar to previous reports (Bauer et al., 2006 and Trenner et al., 2008). The relatively weak and peripheral visual stimulus induced an earlier enhanced activity in low frequencies, a weak enhanced gamma-band activity, and a later reduced high-gamma band activity. When one visual stimulus was paired with two tactile stimuli, subjects often experienced the double-flash illusion (DFI). In trials with the DFI, occipital sensors showed enhanced gamma-band activity. Also, during the DFI, somatosensory sensors showed enhanced low-frequency activity and a reduced beta rebound. The DFI-related enhancements in visual gamma- and somatosensory low-frequency activity were correlated across trials.

Recently, Mishra et al. (2010) have shown that directing attention to the visual stimulus increases the likelihood of perceiving the DFI. Based on this study one might speculate that also spontaneous fluctuations of ongoing attentional processes might influence the perception of the DFI-effect. Our results are in line with this hypothesis: attention has been found to (1) reduce the beta rebound in somatosensory cortex (Bauer et al., 2006 and Trenner et al., 2008) and (2) enhance gamma-band activity in visual cortex (e.g. Fries et al., 2008). This interpretation proposes spontaneous fluctuations of attention as the underlying cause of both somatosensory and visual effects. In support of this, we found the DFI-related enhancements in somatosensory low-frequency and visual gamma-band activity to be correlated across trials. Notably, states of putatively enhanced attention were not always related to more veridical perception, in agreement with previous studies (Yeshurun and Carrasco, 1998).

We find DFI effects and effects to purely visual stimulation in occipital sensors at rather high frequencies. Previous studies using MEG have reported initial high gamma response to visual stimulation before the response settles to a lower and sustained frequency band (Hoogenboom et al., 2006). Since our stimuli were presented for a short duration (16 ms), the observed effects at high frequencies might reflect the previously reported initial high gamma response before the responses settles to lower frequencies. Given the rather long latency of the somatosensory DFI effect in the beta-band, it seems unlikely that it is purely driven by bottom-up stimulus processing. Rather, it might be related to stimulus replay or other top-down mechanisms. For four reasons, however, it is unlikely that this effect is related to motor responses or response preparation: First, the effect is observed contralateral to stimulus presentation but ipsilateral to the response hand. An effect related to response preparation should have a spatial maximum contralateral to the response hand. By contrast, the observed maximum contralateral to stimulation suggests an effect at the sensory stage. Second, a control analysis of rhythmic activity contralateral to the response hand (mirroring the somatosensory ROI across the midline), did not reveal any significant effects (data not shown). Third, if the DFI effect had been due to response preparation then it should also have been visible in
the analysis of the other conditions for which subjects sometimes reported one and sometimes two stimuli (i.e. conditions v2t0 and v2t1), because the response patterns (i.e. button presses) for these conditions and DFI trials were similar. However, the analysis of these conditions did not reveal any significant effects over somatosensory cortex. Finally, the response buttons relevant for the DFI contrast (one and two stimuli) were counterbalanced across subjects. Therefore, it is unlikely that the effects were due to different preparatory signals to different fingers.

One potential concern in the interpretation of neuronal correlates of the DFI is that the DFI might be based on an ambiguity that might in principle as well be related to perception as to decision. Signal detection theory provides a measure (d') of the difference in subjects' perceptual representations between two conditions, independent of the subjects' response bias. In agreement with previous studies on both the audio-visual and visuo-tactile DFI, our d'-analysis revealed that the DFI cannot be explained solely as a consequence of shift in response bias (McCormick & Mamassian, 2008, Mishra et al., 2007, Violentyev et al., 2005 and Watkins et al., 2006). In addition, the perceptual nature of the phenomenon is clearly favored by the phenomenological experience during the illusion. This is supported by the confidence ratings of our subjects, which were very high throughout, in agreement with a recent study addressing this issue in detail (McCormick and Mamassian, 2008).

Several previous studies have investigated the role of rhythmic neuronal activity in cross-modal integration, and to the best of our knowledge, they all investigated the integration of a visual stimulus with an auditory one. Some of those studies compared unimodal visual stimulation with bimodal stimulation and reported supra-additive responses in the bimodal condition (Sakowitz et al., 2001 and Senkowski et al., 2007). Two EEG studies used the audio-visual DFI and found that the illusion is related to an enhanced activity in a lower gamma-band between 30 and 50 Hz in occipital electrodes (Bhattacharya et al., 2002 and Mishra et al., 2007). The timing of these effects was very similar to the high-gamma DFI effect described here. By contrast, the frequency range was lower and this could be due to differences between EEG and MEG, between the analyses, and/or between the paradigms. Any of these differences might also explain why these previous studies did not report two of our central findings: (1) The similarity between the DFI effect and the tactile stimulation effect in the high gamma-frequency range; (2) DFI effects on rhythmic activity over somatosensory areas, or more generally the non-visual areas involved in the induction of the illusion.

Although Mishra et al. (2007) did not report any changes in rhythmic activity over non-visual (i.e. auditory) areas, they found an enhanced evoked component in the ERPs in sensors over auditory cortex ranging from 92 to 124 ms if subjects perceive the illusion. For DFI trials, we found a significant enhancement of low-frequencies over somatosensory areas. While the timing suggests that this effect might be related to the effect described by Mishra et al. for the audio-visual DFI, we did not find any significant contribution of stimulus-locked effects.

Our results suggest that the DFI occurred during states of enhanced attention and previous studies would suggest that in this case, the DFI should also have been associated with enhanced somatosensory gamma-band activity (Bauer et al., 2006). While somatosensory gamma-band activity was indeed enhanced at a similar time range as described by Bauer et al., this did not reach statistical significance (Fig. 5A), probably due to insufficient signal-to-noise-ratio or statistical power.
However, we did find that the DFI is related to significantly enhanced gamma-band activity over visual areas. This DFI related gamma-band activity over visual areas had a strong similarity to the tactile stimulation induced gamma-band activity over somatosensory areas and it occurred ~ 100-200 ms earlier in time and topographically more central than the gamma-band increase in response to isolated visual stimulation. We therefore propose that the DFI is related to a somatosensory-to-visual transmission of gamma-band activity, which occurs more likely during spontaneously enhanced attention. The consequently enhanced gamma-band activity in visual cortex during DFI will most likely enhance the impact of visual cortex on other brain areas.

It should be noted that in this hypothesis somatosensory gamma-band activity does not need to be enhanced. Rather, we speculate the DFI occurs when a given amount of somatosensory gamma-band activity is more effectively transmitted to visual cortex. While the interpretation of a transmission of gamma-band activity from somatosensory to visual cortex remains speculative, there are several studies that might propose a framework for this hypothesis. A functional interaction of somatosensory and visual cortex has been reported previously. For example, studies using transcranial magnetic stimulation provided evidence for a functional role of visual cortex for tactile discrimination tasks (Merabet et al., 2004 and Zangaladze et al., 1999). Macaluso et al. (2002) reported an increased BOLD response in visual cortex if visual stimuli are accompanied by spatially congruent tactile stimuli. Also, ERP studies reported a functional modulation of visual cortex by attention to tactile stimuli (Eimer & Driver, 2000 and Eimer & Van Velzen, 2002). Recently, it has been shown that selective spatial attention to tactile stimuli modulates rhythmic activity in low frequencies in occipital areas (Bauer et al., 2006).

We can only speculate about the anatomical framework that might underlie the transmission of rhythmic activity from somatosensory areas to visual areas. While there is evidence for direct connections between early visual cortex and auditory cortex (Bizley et al., 2007, Falchier et al., 2002 and Rockland & Ojima, 2003) and studies indicate the possibility of direct connections of auditory and somatosensory cortices (Fu et al., 2003 and Hackett et al., 2007), there is to our knowledge no evidence for similar connections from somatosensory cortex to visual cortex. Alternatively, the transmission might be relayed by higher-level multimodal areas (e.g. the superior temporal sulcus, STS) which control information transfer between sensory cortices. Such a mechanism has been suggested in fMRI studies on visuo-tactile attention tasks (Macaluso et al., 2000) as well as on audio-visual integration (Noesselt et al., 2007). In line with this hypothesis, a recent study found interactions of neurons in the STS and auditory cortex for face-voice integration (Ghazanfar et al., 2008). A third possible mechanism might be an information transfer via subcortical areas. Several studies found multimodal thalamo-cortical projections in rats (e.g. Barth et al., 1995 and Nicolelis et al., 1997) and recent studies on monkeys have suggested a role of the thalamic system in audio-tactile integration by rhythmic neural activity (Lakatos et al., 2007).

Lakatos and colleagues found in their study that somatosensory inputs modulated auditory responses by changing the phase of ongoing auditory cortical oscillations. Low frequent rhythmic activity was phase reset by somatosensory stimulation which subsequently enhanced auditory induced responses when auditory inputs fell on peaks of low-frequency activity and vice versa. A similar mechanism
might account for the findings in our study. However, due to the technical limitations of our method, this remains speculative.

In summary, our results elucidate some of the mechanisms behind the double flash illusion. They add to the growing evidence for a functional role of rhythmic neuronal synchronization in cognition. We hypothesize that the double flash illusion occurs when tactile induced gamma-band activity is transmitted to visual cortex. During ongoing fluctuations in attention, this transmission is more likely during patterns of rhythmic activity associated with states of enhanced attention. Further studies will be needed to confirm this hypothesis.

Acknowledgments
We are most grateful to Sander Berends for collecting parts of the MEG data. We also thank Pieter Buur and Paul Gaalman for help with MRI recordings. JL is supported by the German Research Foundation (LA 2400/1-1), PF is supported by the Human Frontier Science Program Organization, The Netherlands Organization for Scientific Research, The Volkswagen Foundation, and the European Science Foundation's European Young Investigator Award Program.

References


The following are the supplementary materials related to this article.

Figure S1 Time–frequency power spectra for the condition v1t2 in visual sensors and high frequencies, sorted for subjects' visual perception. (A) Trials in which subjects perceived one visual stimulus, (B) trials in which subjects reported two visual stimuli. The dashed box indicates the time–frequency range used for statistical analysis of differences between both conditions. The colorbar applies to both plots.

Figure S2 Analysis of eye-movements: The double flash contrast in EOG-sensors. The analysis was performed for the stimulation condition v1t2. Positive/negative t-values indicate stronger/weaker eye movements in trials in which subjects perceived an illusory second flash as compared to trials in which they perceived the veridical single flash (see Methods for details). The statistical comparison was done as for the MEG signals, but no significant clusters were found.

Figure S3 Statistical analysis of the stimulus-locked activity in DFI trials in somatosensory (A) and visual (B) sensors in the time-domain (event-related fields). Positive t-values indicate higher field amplitudes in trials in which subjects perceived an illusory second flash as compared to trials in which they perceived the veridical single flash. No significant clusters were found.
Perception of the touch-induced visual double-flash illusion correlates with changes of rhythmic neuronal activity in human visual and somatosensory areas. NeuroImage 54(2), 1395-1405. doi: http://dx.doi.org/10.1016/j.neuroimage.2010.09.031

Figure S4 Statistical analysis of the stimulus-locked activity in DFI trials in somatosensory sensors for spectral power. Positive t-values indicate greater power in trials in which subjects perceived an illusory second flash as compared to trials in which they perceived the veridical single flash. No significant clusters were found.

Figure S5 Statistical analysis of the stimulus-locked activity in DFI trials in visual sensors for spectral. Positive t-values indicate greater power in trials in which subjects perceived an illusory second flash as compared to trials in which they perceived the veridical single flash. No significant clusters were found.