



## Electrical neuroimaging of memory discrimination based on single-trial multisensory learning

Antonia Thelen<sup>a</sup>, Céline Cappe<sup>b</sup>, Micah M. Murray<sup>a,c,d,e,\*</sup>

<sup>a</sup> The Functional Electrical Neuroimaging Laboratory, Department of Clinical Neurosciences, Vaudois University Hospital Center and University of Lausanne, Lausanne, Switzerland

<sup>b</sup> Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

<sup>c</sup> Radiology Department, Vaudois University Hospital Center and University of Lausanne, Lausanne, Switzerland

<sup>d</sup> Electroencephalography Brain Mapping Core, Center for Biomedical Imaging (CIBM) of Lausanne and Geneva, Switzerland

<sup>e</sup> Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

### ARTICLE INFO

#### Article history:

Accepted 10 May 2012

Available online 18 May 2012

#### Keywords:

Multisensory

Cross-modal

Event-related potential (ERP)

Memory

Object recognition

Auditory cortex

### ABSTRACT

Multisensory experiences influence subsequent memory performance and brain responses. Studies have thus far concentrated on semantically congruent pairings, leaving unresolved the influence of stimulus pairing and memory sub-types. Here, we paired images with unique, meaningless sounds during a continuous recognition task to determine if purely episodic, single-trial multisensory experiences can incidentally impact subsequent visual object discrimination. Psychophysics and electrical neuroimaging analyses of visual evoked potentials (VEPs) compared responses to repeated images either paired or not with a meaningless sound during initial encounters. Recognition accuracy was significantly impaired for images initially presented as multisensory pairs and could not be explained in terms of differential attention or transfer of effects from encoding to retrieval. VEP modulations occurred at 100–130 ms and 270–310 ms and stemmed from topographic differences indicative of network configuration changes within the brain. Distributed source estimations localized the earlier effect to regions of the right posterior temporal gyrus (STG) and the later effect to regions of the middle temporal gyrus (MTG). Responses in these regions were stronger for images previously encountered as multisensory pairs. Only the later effect correlated with performance such that greater MTG activity in response to repeated visual stimuli was linked with greater performance decrements. The present findings suggest that brain networks involved in this discrimination may critically depend on whether multisensory events facilitate or impair later visual memory performance. More generally, the data support models whereby effects of multisensory interactions persist to incidentally affect subsequent behavior as well as visual processing during its initial stages.

© 2012 Elsevier Inc. All rights reserved.

### Introduction

Perceptions are often affected by the way in which information from several sensory modalities are combined, i.e. under multisensory conditions (Calvert, 2004; Stein and Meredith, 1993). In addition to their immediate effects on perception, multisensory experiences at one point in time can impact unisensory processing during later encounters (Gottfried et al., 2004; Kim et al., 2008; Lehmann and Murray, 2005; Meylan and Murray, 2007; Murray et al., 2004, 2005; Shams and Seitz, 2008). The circumstances under which such multisensory memory traces impact subsequent unisensory retrieval remain unresolved and were the focus of the present study. This issue falls within the larger framework of differential mechanisms and consequences of multisensory versus unisensory learning (Murray and Sperdin, 2010; Shams and Seitz, 2008; Shams et al., 2011).

The impact of multisensory learning upon unisensory (visual) recognition has been studied mainly through two paradigms. In one paradigm, subjects underwent extensive multisensory training before unisensory recognition performance was assessed in a separate retrieval session (Nyberg et al., 2000; Wheeler and Petersen, 2000). The task was to explicitly remember the context in which stimuli had been presented during encoding (unisensory or multisensory). In another paradigm, the impact of multisensory experiences upon subsequent unisensory recognition was assessed through single-trial learning during a continuous recognition task (Lehmann and Murray, 2005; Murray et al., 2004, 2005). Encoding and retrieval were separated by (relatively) short time intervals in the same experimental block, and subjects indicated whether or not images were being presented for the first or repeated time.

These paradigms have led to discordant results both in terms of performance and brain activity. For example, Nyberg et al. (2000) observed relatively poorer memory performance accuracy for words that had been paired with sounds during encoding vs. words that had not (76% vs. 84%, respectively). In terms of brain activity during

\* Corresponding author at: Department of Clinical Neurosciences, CHUV, BH08.078, Rue du Bugnon 46, 1011 Lausanne, Switzerland. Fax: +41 21 314 1319.

E-mail address: [Micah.Murray@chuv.ch](mailto:Micah.Murray@chuv.ch) (M.M. Murray).

the retrieval phase, these authors found that auditory cortices were active in response to visually presented words that had been paired with corresponding sounds during the encoding phase (see also Wheeler and Petersen, 2000 for similar findings with pictures of objects). They took this activity pattern as evidence that networks active during encoding are re-activated during retrieval; a pattern more generally consistent with theories of redintegration (Hamilton, 1859). Under the framework of redintegration, a component part of a consolidated memory is sufficient to (re)activate the whole experience's representation. If the consolidated memory, for example, entails both auditory and visual components, then stimulating with sounds would lead to both auditory and visual cortical activity (see also Rugg et al., 2008 for a similar framework).

By contrast, our group has consistently observed improved memory performance for images that had been paired with their corresponding sounds versus those images only ever presented in a unisensory manner (Lehmann and Murray, 2005; Murray et al., 2004, 2005; Murray and Sperdin, 2010). Plus, both electrical and hemodynamic imaging studies indicated that regions of the lateral occipital cortices differentiate between image repetitions that had been previously paired with sounds or not (Murray et al., 2004, 2005). Stronger responses were observed for visual stimuli that had been previously presented with a semantically congruent sound. There was no evidence for effects within auditory cortices in these studies by Murray and colleagues. Moreover, the timing of effects observed in Murray et al. (2004) would suggest that past multisensory experiences impact the initial stages of visual processing (i.e. within the initial 100 ms post-stimulus onset). Effects within visual cortices do not appear to be limited to tasks requiring visual discrimination. For example, von Kriegstein and Giraud (2006) observed activation changes within the fusiform face area in response to voices as a function of prior learned associations with faces. Others have observed activations within visual cortices in response to meaningless complex sounds previously paired with meaningless images (though activations were smaller than in response to visual stimuli) (Butler and James, 2011). In this study, the magnitude of the activation did not appear to be directly linked to subjects' accuracy in indicating if a given sound/image had been previously paired with a sound vs. image.

These discrepant patterns of brain activity (i.e. differential effects within visual vs. auditory cortices) may stem from paradigmatic differences, including whether or not subjects must explicitly discriminate the context in which a given stimulus had been paired. Another alternative is that the activation of auditory regions in response to visual stimuli during memory-related tasks is a hallmark of performance impairments rather than redintegration per se. A continuous recognition task has been used to identify conditions under which memory performance is impaired for image repetitions that were initially presented in a multisensory context. While performance was enhanced for stimuli that had been initially encountered in a semantically congruent multisensory context, it was unchanged if initially encountered in a semantically incongruent context, and was impaired if initially encountered with a meaningless sound (i.e. pure tone) (Lehmann and Murray, 2005). All of these effects were relative to performance with images that were initially encountered in a unisensory context (i.e. the same image repeated once). It is likewise important to note that these effects on accuracy did not co-occur with parallel modulations in reaction times (in fact, reaction times did not differ) and were not the consequence of carry-over effects from encoding (reviewed in Murray and Sperdin, 2010; see also Baier et al., 2006 for similar work on multisensory expectancy). The performance impairment we observed with meaningless sounds was confounded by the fact that the same sound was paired with multiple objects across trials, thereby making it possible that the representation of any given visual object is effectively contaminated by a constant distracter (here the auditory stimulus).

The present study sought to resolve the role of episodic pairings on single-trial multisensory memory processes by combining psychophysics with electrical neuroimaging of VEPs. On the one hand, if single-trial episodic events are sufficient to generate a perceptual/memory trace then performance accuracy would be predicted to be impaired for repetitions of images that had been paired with sounds on their initial encounter. This would suggest that the observations of Lehmann and Murray (2005) are due to the episodic pairing rather than to the repeated association/dissociation of the same sound with multiple images across trials. Alternatively, no performance differences (vs. images that had only been encountered visually) would be predicted if the unique, meaningless sounds are treated as if they were an incongruent event akin to pairing the sound of one (known) object with the image of another object. On the other hand, prior brain imaging data would suggest that differential responses to incoming visual stimuli can manifest as early as ~100 ms post-stimulus onset and are indicative of the incidental discrimination of stored object representations according to past (semantic) multisensory experiences (Murray et al., 2004). To the extent that single-trial episodic pairings are similarly effective in establishing distinct object representations, we should observe differential VEPs at equally early latencies. Because electrical neuroimaging analyses distinguish between effects due to modulations in response strength and response topography (the latter of which is forcibly indicative of generator changes), we could likewise assess if and when distinct generator configurations respond to visual stimuli previously encountered in a unisensory vs. multisensory manner. When combined with distributed source estimations, we can then disentangle whether or not differential activity, putatively arising within auditory cortices and perhaps elsewhere, is linked to performance decrements or to redintegration processes.

## Materials and methods

### Participants

The experiment included 22 (11 women) volunteers aged between 23 and 30 years (mean age  $\pm$  SD = 26.45  $\pm$  3.1 years). The last 12 of these completed the psychophysics paradigm while EEG was simultaneously recorded (detailed below). The study was conducted in accordance with the Declaration of Helsinki, and all subjects provided their written informed consent to participate in the study. The experimental procedures were approved by the Ethics Committee of the Vaudois University Hospital Center and University of Lausanne. Nineteen subjects were right-handed and the remaining 3 left-handed, according to the Edinburgh Inventory (Oldfield, 1971). No subject had a history of neurological or psychiatric illness, and all subjects had normal or corrected-to-normal vision as well as reported normal hearing.

### Task

Subjects performed a continuous recognition task, which required the discrimination of initial from repeated presentations of line drawings that were pseudo-randomized within a block of trials. They were instructed to perform as quickly and as accurately as possible. Further, each object (irrespective of whether it was initially presented in a unisensory or multisensory context) was only repeated once throughout the duration of the experiment.

The pictures were subdivided into two groups. Initial presentations were either unisensory or multisensory. Repeated presentations were always unisensory. Thus, half of the repeated presentations had been multisensory when initially encountered and the other half had been unisensory when initially encountered. We will refer to our experimental conditions as V, for the visual-only initial condition and V— for the same stimuli when repeated. AV will refer to the initial

auditory-visual presentation, whereas V+ will refer to the repeated presentation of the visual component of these stimuli (Fig. 1a).

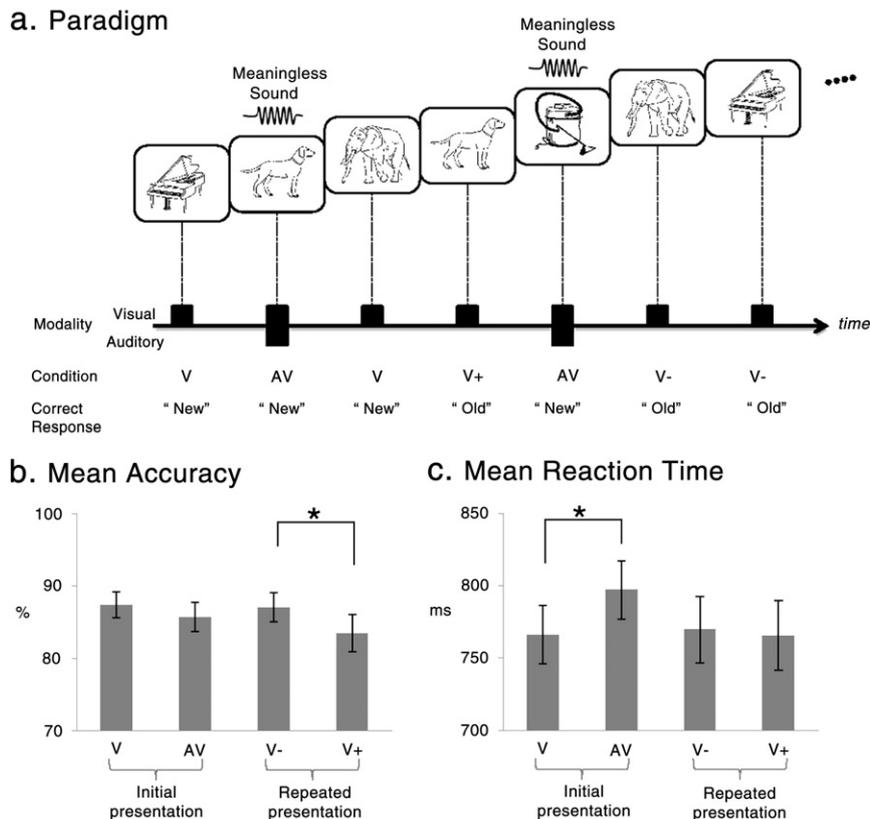
The line drawings were taken from a standardized set (Snodgrass and Vanderwart, 1980) or obtained from an online library (dgl.microsoft.com), and included a mix of living and non-living stimuli (see Appendix for full list). The pictures were equally subdivided over experimental conditions and blocks as described above. Plus, the different categories of objects were equally intermixed. In order to minimize the possibility that the observed effects were due to low-level visual features, we analyzed the spatial frequency spectra and the luminance between the two image groups (AV and V), and we did not find any difference between image groups. The full details of these procedures have been reported by Knebel et al. (2008). The images were presented centrally and appeared black on a white background. On initial presentations these visual stimuli could (50%) or could not be paired with a meaningless sound (created with Adobe Audition 1.0). These sounds differed in their spectral composition, ranging from 100 Hz to 4700 Hz and were sometimes modulated in terms of amplitude envelopes and/or waveform types (triangular and sinusoid). All sounds were 500 ms duration (10 ms rise/fall, in order to avoid clicks; 16bit stereo; 44100 Hz digitization).

All stimuli were presented for 500 ms, followed by a randomized inter-stimulus interval (ISI) ranging from 900 to 1500 ms. The mean ( $\pm$ SD) number of trials between the initial and the repeated presentation of the same image was  $9 \pm 4$  pictures for either presentation condition (V and AV). Also the distribution of old and new pictures throughout the length of the block was controlled, so as to avoid response-decision bias and to maintain an equal probability of a “new” object across quartiles within a block. This type of bias refers to subjects being able to calculate predictive probabilities about the

upcoming stimuli and responses, which could lead to faster reaction times and/or a drop in attention. Within a block there were 136 trials, equally divided between V, AV, V–, and V+ conditions (i.e. 34 trials each). This is identical to the block length used in our prior studies (Lehmann and Murray, 2005; Murray et al., 2004).

The experiment took place in sound-attenuated chamber, where subjects were seated centrally in front of a 20” computer monitor (HP LP2068), located about 140 cm away from them (visual angle  $\sim 4^\circ$ ). The auditory stimuli were presented over insert earphones (Etymotic model: ER4S), and the volume was adjusted to a comfortable level ( $\sim 62$  dB). The stimuli were all presented and controlled by E-Prime 2.0, and all behavioral data were recorded in conjunction with the serial response box (Psychology Software Tools, Inc.; [www.pstnet.com](http://www.pstnet.com)). All participants completed 2 blocks of trials. While this paradigm indeed introduces a degree of spatial disparity between the auditory and visual modalities (when stimuli are multisensory), we do not believe that such would be impeding any interactions here. Prior research has reliably documented integrative effects on behavior and brain responses using similar setups (e.g. Cappe et al., 2010, 2012; Raji et al., 2010).

The behavioral data were treated as follows: Mean RT (in [ms]) and accuracy (%) of correct responses were calculated for each subject and condition (V, AV, V–, and V+) separately. We then performed paired *t*-tests, specifically comparing the initial encoding conditions (V vs. AV) or the repeated retrieval conditions (V– vs. V+). Finally, we sought to assess whether effects during encoding (i.e. differences between V vs. AV) were directly linked to and/or predictive of effects during retrieval (i.e. differences between V– vs. V+), given that prior studies would suggest that these are dissociable (reviewed in Murray and Sperdin, 2010). As will be made clearer in the Results below, we



**Fig. 1.** (a) Illustration of the paradigm used to investigate multisensory encoding and recall in working memory. (b) Group-averaged accuracy ( $\pm$  s.e.m.) for the four experimental conditions: initial presentations of visual (V) and auditory-visual (AV) and their repeated presentation (V– and V+, respectively). (c) Group-averaged mean reaction times ( $\pm$  s.e.m.) for the same experimental conditions as in (b). Significant effects ( $p < 0.05$ ) are marked with an asterisk.

calculated the correlation coefficient between the differences in reaction time during encoding and the difference in performance accuracy during retrieval.

#### EEG acquisition and pre-processing

Continuous EEG was acquired from 160 scalp electrodes (sampling rate at 1024 Hz) using a Biosemi ActiveTwo system from a subset of 12 subjects (3 women; mean age  $\pm$  SD = 27.1  $\pm$  3.5 years) who were the last of the abovementioned 22 participants to take part in the psychophysics paradigm. Data pre-processing and analyses were performed using Cartool (Brunet et al., 2011; <http://sites.google.com/site/fbmlab/cartool>). Epochs from 100 ms pre-stimulus to 500 ms post-stimulus onset were averaged for each of the four experimental conditions and from each subject in order to calculate the VEPs. In addition to a  $\pm$  80  $\mu$ V artifact rejection, EEG epochs containing eye blinks or other noise transients were removed based on a trial-by-trial visual inspection of the data. Before group averaging, data from artifact electrodes of each subject were interpolated using 3-D splines (Perrin et al., 1987). On average, 5 of the 160 channels were interpolated (range 2–12). ERP data were baseline corrected using the pre-stimulus period, band-pass filtered (0.1–60 Hz including a notch at 50 Hz) and recalculated against the average reference. On average, there were 60 (range 43–68) and 61 (range 42–68) accepted epochs for the V– and V+ conditions, respectively.

#### VEP analyses

The VEP analyses were based on the hypothesis that a differential neural response would be found between the V+ and V– conditions (Murray et al., 2004, 2005). The approach we used here has been referred to as electrical neuroimaging and is based largely on the multivariate analysis of global features of the electric field at the scalp that in turn inform the selection of time periods for analyses of source estimations (Michel and Murray, 2012; Murray et al., 2008; Tzovara et al., in press). These electrical neuroimaging analyses allowed us to differentiate effects following from modulations in the strength of responses of statistically indistinguishable brain generators from alterations in the configuration of these generators (viz. the topography of the electric field at the scalp), as well as latency shifts in brain processes across experimental conditions. Additionally, we applied the local auto-regressive average distributed linear inverse solution (LAURA; Grave de Peralta Menendez et al., 2001, 2004) to visualize and statistically contrast the likely underlying sources of effects identified during the preceding analysis steps of the surface-recorded VEPs.

As a first level of analysis, we analyzed waveform data from all electrodes as a function of time post-stimulus onset in a series of pair-wise comparisons (*t*-tests) between responses to the V+ and V– conditions. Temporal auto-correlation at individual electrodes was corrected through the application of an 11 contiguous data-point temporal criterion ( $\sim$  10 ms at 1024 Hz sampling) for the persistence of differential effects (Guthrie and Buchwald, 1991). Similarly, spatial correlation was addressed by considering as reliable only those effects that entailed at least 5% of the electrodes from the 160-channel montage (i.e. 8). These combined criteria were applied to correct for multiple comparisons, though we note that our conclusions are based on reference-independent global measures of the electric field at the scalp. We would therefore also emphasize that the number of electrodes exhibiting an effect at a given latency will depend on the reference, and this number is not constant across choices of reference because significant effects are not simply re-distributed across the montage (discussed in Tzovara et al., in press). Likewise, the use of an average reference receives support from biophysical laws as well as the implicit re-centering of VEP data to such when performing source estimations (discussed in Brunet et al., 2011). Analyses of VEP voltage waveform data (vs. the average reference) are presented here to provide a clearer

link between canonical VEP analysis approaches and electrical neuroimaging. The results of this VEP waveform analysis are presented as an area plot that shows the number of electrodes exhibiting a significant effect as a function of time (relative to stimulus onset). This type of display was chosen to provide a sense of the dynamics of a statistical effect between conditions. While these analyses give a visual impression of specific effects within the dataset, our conclusions are principally based on reference-independent global measures of the electric field at the scalp that are described below.

The global electric field strength was quantified using global field power (GFP) (Lehmann and Skrandies, 1980). This measure is equivalent to the standard deviation of the voltage potential values across the entire electrode montage at a given time point and represents a reference-independent measure of the VEP strength (Murray et al., 2008; Koenig and Melie-Garcia, 2010). GFP was statistically contrasted using a millisecond-by-millisecond paired *t*-test in conjunction with the abovementioned temporal criterion for significant effects to correct for multiple contrasts. While this dependent measure provides an assay of VEP strength, it is inherently insensitive to spatial (i.e. topographic) variation in the VEP across conditions.

In order to test the VEP topography independently of its strength, we used Global Dissimilarity (DISS) (Lehmann and Skrandies, 1980). DISS is equivalent to the square root of the mean of the squared difference between the potentials measured at each electrode for different conditions, normalized by the instantaneous GFP. It is also directly related to the (spatial) correlation between two normalized vectors (cf. Appendix in Murray et al., 2008). We then performed a non-parametric randomization test (TANOVA, Murray et al., 2008). The DISS value at each time point is compared to an empirical distribution derived from permuting the condition label of the data from each subject. Because changes in topography forcibly follow from changes in the configuration of the underlying active sources (Lehmann, 1987), this analysis reveals when the experimental conditions activated distinct sets of brain networks.

The collective post-stimulus group-average VEPs were subjected to a topographic cluster analysis based on a hierarchical clustering algorithm (Murray et al., 2008). This clustering identifies stable electric field topographies (hereafter template maps). The clustering is exclusively sensitive to topographic modulations, because the data are first normalized by their instantaneous GFP. The optimal number of temporally stable VEP clusters (i.e. the minimal number of maps that accounts for the greatest variance of the dataset) was determined using a modified Krzanowski-Lai criterion (Murray et al., 2008). The clustering makes no assumption on the orthogonality of the derived template maps (De Lucia et al., 2010a, 2010b; Pourtois et al., 2008). Template maps identified in the group-average VEP were then submitted to a fitting procedure wherein each time point of each single-subject VEP is labeled according to the template map with which it best correlated spatially (Murray et al., 2008) so as to statistically test the relative presence of each template map in the moment-by-moment scalp topography of the VEP and the differences in such across conditions. These values can be expressed as the probability of a given template map yielding a higher spatial correlation in the single-subject data from each condition. Statistical analysis of these values was performed with an analysis of variance (ANOVA).

#### Source estimations

We estimated the localization of the electrical activity in the brain using a distributed linear inverse solution (minimum norm) applying the LAURA regularization approach comprising biophysical laws as constraints (Grave de Peralta Menendez et al., 2001, 2004; see also Michel et al., 2004 for review). LAURA selects the source configuration that better mimics the biophysical behavior of electric vector fields (i.e. activity at one point depends on the activity at neighboring points according to electromagnetic laws). In our study, homogenous

regression coefficients in all directions and within the whole solution space were used. LAURA uses a realistic head model, and the solution space included 4024 nodes, selected from a  $6 \times 6 \times 6$  mm grid equally distributed within the gray matter of the Montreal Neurological Institute's average brain (courtesy of R. Grave de Peralta and S. Gonzalez Andino; <http://www.electrical-neuroimaging.ch/>). Prior basic and clinical research from members of our group and others has documented and discussed in detail the spatial accuracy of the inverse solution model used here (e.g. Gonzalez Andino et al., 2005; Grave de Peralta Menendez et al., 2004; Martuzzi et al., 2009; Michel et al., 2004). In general, the localization accuracy is considered to be along the lines of the matrix grid size (here 6 mm). The results of the above topographic pattern analysis defined time periods for which intracranial sources were estimated and statistically compared between conditions (here 73–113 ms post-stimulus). Prior to calculation of the inverse solution, the VEP data were down-sampled and affine-transformed to a common 111-channel montage. Statistical analyses of source estimations were performed by first averaging the VEP data across time to generate a single data point for each participant and condition. This procedure increases the signal-to-noise ratio of the data from each participant. The inverse solution was then estimated for each of the 4024 nodes. These data were then submitted to a paired *t*-test.

We combined two statistical criteria for concluding that an effect was reliable. First, a spatial extent criterion of at least 17 contiguous significant nodes was applied (see also Cappe et al., 2010, 2012; De Lucia et al., 2010a, 2010b; Knebel and Murray, 2012; Knebel et al., 2011; Toepel et al., 2009 for a similar spatial criterion). This spatial criterion was determined using the AlphaSim program (available at <http://afni.nimh.nih.gov>) and assuming a spatial smoothing of 6 mm full-width half maximum. This criterion indicates that there is a 3.54% probability of a cluster of at least 17 contiguous nodes, which gives an equivalent node-level *p*-value of  $p \leq 0.0002$ . Second and because distributed source models yield non-zero values in all solution points, it is conceivable that statistical effects will be obtained in nodes that are weakly responsive (i.e. have current density values close to zero or alternatively well below the mean across the entire set of nodes in the brain volume). To minimize the contribution of such “erroneous” or “ghost” sources, we removed all nodes with current density values less than or equal to two standard deviations below the volume's mean within each condition (here,  $V+$ : mean  $\pm$  SD =  $0.0073 \pm 0.0025 \mu\text{A}/\text{mm}^3$  and  $V-$ : =  $0.0071 \pm 0.0026 \mu\text{A}/\text{mm}^3$ ). In this way, we sought to limit statistical effects to nodes that could reasonably be described as “active” sources. The results of the source estimations were rendered on the Montreal Neurologic Institute's average brain with the Talairach and Tournoux (1988) coordinates of the largest statistical differences within a cluster indicated.

### Correlation analysis

To test whether there was a linear relationship between brain activity within source estimations and behavior, we correlated (Pearson's correlation after first testing for normality of the distributions using a Kolmogorov–Smirnov test) the difference in performance accuracy between  $V+$  and  $V-$  conditions with the corresponding difference in source strength. Specifically, we subtracted the percent correct performance on the  $V+$  condition from that from the  $V-$  condition for each subject. The  $[(V+) - (V-)]$  difference was calculated for the mean activity within the significant clusters (pSTG/STS and mSTG/MTG; see Results for details) for every subject.

## Results

### Behavioral data

Analysis of performance accuracy (Fig. 1b) revealed that participants were equally capable of indicating the initial presentation of

stimuli both when presented in a multisensory auditory-visual context ( $AV = 85.7 \pm 2\%$ ; mean  $\pm$  s.e.m.) and also when presented in a unisensory visual context ( $V = 87.4 \pm 1.8\%$ ;  $t_{(21)} = 1.452$ ;  $p = 0.161$ ;  $\eta_p^2 = 0.681$ ). There was therefore no indication of encoding differences in terms of discrimination accuracy. By contrast, performance when indicating image repetitions was significantly impaired when the initial presentation had entailed a multisensory vs. unisensory context (i.e.  $V+$  vs.  $V-$ ;  $83.5 \pm 2.6\%$  vs.  $87 \pm 2\%$ ;  $t_{(21)} = -2.38$ ;  $p = 0.027$ ;  $\eta_p^2 = 0.674$ ). This difference indicates that image repetitions are incidentally discriminated according to past multisensory experiences. Though not a priori part of our research aims, we also tested the role of object category (i.e. living vs. man-made) on image repetition discrimination as a function of prior multisensory vs. unisensory pairings. The accuracy data were subjected to a  $2 \times 2$  within subjects ANOVA with factors of object category (living and man-made) and experimental condition ( $V-$  and  $V+$ ). There was a significant main effect of experimental condition ( $F_{(1,21)} = 4.522$ ;  $p = 0.045$ ;  $\eta_p^2 = 0.177$ ). Neither the main effect of object category ( $p = 0.058$ ) nor the interaction ( $p = 0.797$ ) reached the 0.05 significance criterion.

In contrast with the pattern observed with accuracy rates (but highly consistent with prior work; e.g. Lehmann and Murray, 2005), mean reaction times differed for initial presentations, but not for image repetitions (Fig. 1c). Specifically, reaction times to initial presentations under multisensory conditions were significantly slower than those to initial presentations under unisensory conditions ( $AV$  vs.  $V = 797 \pm 20$  ms vs.  $766 \pm 20$  ms;  $t_{(21)} = -4.233$ ;  $p < 0.001$ ;  $\eta_p^2 = 0.873$ ). Reaction times to image repetitions did not significantly differ ( $V+$  vs.  $V- = 765 \pm 24$  ms vs.  $769 \pm 23$  ms;  $t_{(21)} = 0.800$ ;  $p = 0.433$ ;  $\eta_p^2 = 0.955$ ). Finally, we examined if there was a linear relationship between reaction time differences during encoding and accuracy differences during retrieval. There was no evidence of a significant correlation ( $r_{(20)} = -0.008$ ;  $p > 0.9$ ), providing no evidence for a carry-over effect.

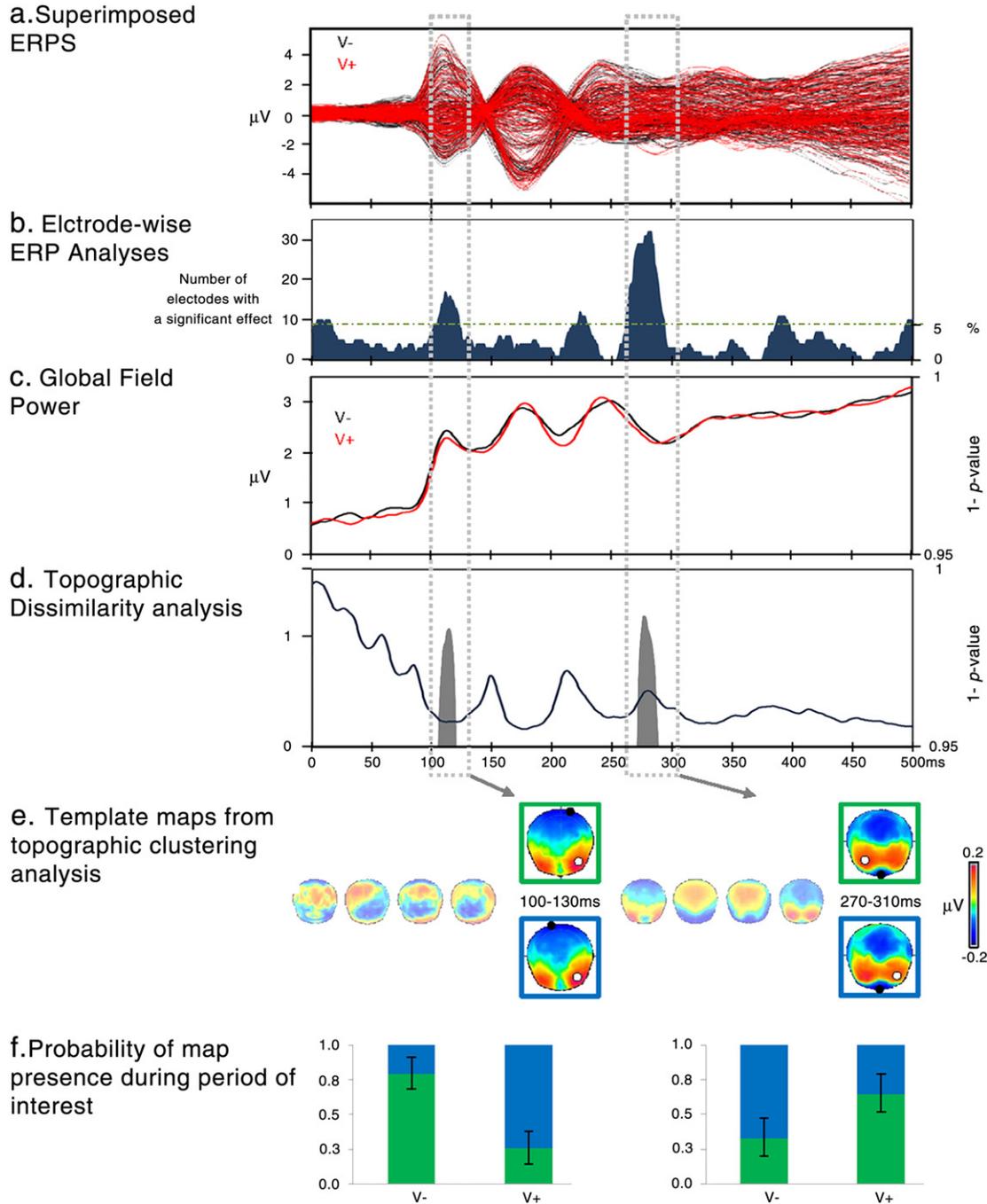
This pattern of results rules out explanations in terms of auditory capture of attention, selective attention, or novel context detection facilitating perceptual memory trace formation (Ranganath and Rainer, 2003), as well as general alerting. If auditory capture of attention were driving our effects, then a significant correlation would have been expected between the magnitude of this capture (quantified as the reaction time difference between visual and multisensory conditions during initial image presentations) and the magnitude of its effect on memory discrimination (i.e. accuracy levels during repeated imaging presentations). Similarly, if selective attention could account for our findings, then faster reaction times would have been expected for initial presentations of multisensory vs. unisensory stimuli (i.e.  $AV$  vs.  $V$ ). Rather, our results suggest that participants were not overtly attending to the auditory channel, which in principle could have cued participants to respond “new.” Slower reaction times for the multisensory condition suggest that participants did not expect auditory events (Spence et al., 2001), and accuracy did not reliably differ. We also rule out possible bias due to general arousal and fatigue, because the relative distribution of “new” and “old” pictures was maintained throughout a block of trials (see Materials and methods).

### Surface VEP data

The VEP analyses focused on differences between the  $V+$  and  $V-$  conditions (see Materials and methods). Fig. 2 displays VEPs from the  $V+$  and  $V-$  conditions (Fig. 2a) as well as the results of electrical neuroimaging analyses (Fig. 2b–f). Analyses of VEP voltage waveforms from the entire electrode montage as a function of time revealed two time periods of significant differences – the first at  $\sim 100$  ms and a second at  $\sim 270$  ms (Fig. 2b) post-stimulus onset. We would remind the reader of the reference-dependent nature of analyses of voltage waveforms. Nonetheless, we supply such here to provide a clearer link between electrical neuroimaging and more

traditional analysis approaches. Fig. 2c displays the group-average GFP waveforms as well as results of *t*-tests on the GFP as a function of time (note that 1 minus *p*-value is displayed). There was no evidence of GFP differences between conditions, providing no evidence for differences in response strength as a function of past multisensory vs. unisensory experience. Fig. 2d displays the DISS between the group-average VEPs from the V+ and V− conditions. Significant topographic differences were observed over the 100–130 ms and 270–

310 ms post-stimulus intervals. We next determined whether these topographic effects stem from the predominance of different stable map configurations in each condition or instead from latency shifts across conditions. We subjected the cumulative group-average VEPs from both conditions to a clustering analysis. This procedure identified 19 template maps that explained 97.26% of variance of the concatenated group-averaged VEP data set. While during most of the post-stimulus period one template map could account for the



**Fig. 2.** Visual evoked potential (VEP) data and results of the stepwise analyses (only  $p < 0.05$  with an 11 consecutive time frame criterion are shown). (a) Group-averaged ( $n = 12$ ) VEP waveforms are shown superimposed across all electrodes for both experimental conditions of repeated image presentations (V− in black and V+ in red). (b) The results of the millisecond-by-millisecond paired *t*-test at each of the scalp electrodes is shown. The total number of electrodes expressing a significant difference at each TF is plotted. The dotted line marks the 5% ( $n = 8$ ) threshold of total electrode montage. (c) GFP analysis across time did not reveal significant modulations between the V− (black trace) and V+ (red trace) conditions. (d) Global dissimilarity analysis revealed two periods of differential responses at 100–130 ms and at 270–310 ms. The dark blue line displays the DISS as a function of time. Significant differences based on a non-parametric permutation test are shown in gray. The scale on the right indicates 1 minus *p*-value (range: 0.95–1). (e) The template maps resulting from the topographic pattern analysis are shown. During most of the post-stimulus period the same template map was observed in the group-average VEPs from both conditions. During two periods, different maps were observed as a function of experimental condition in the group-averaged VEPs. (f) The histograms depict the result of the spatial correlation fitting for both periods. During each period one of the two template maps fitted better to one of the experimental conditions.

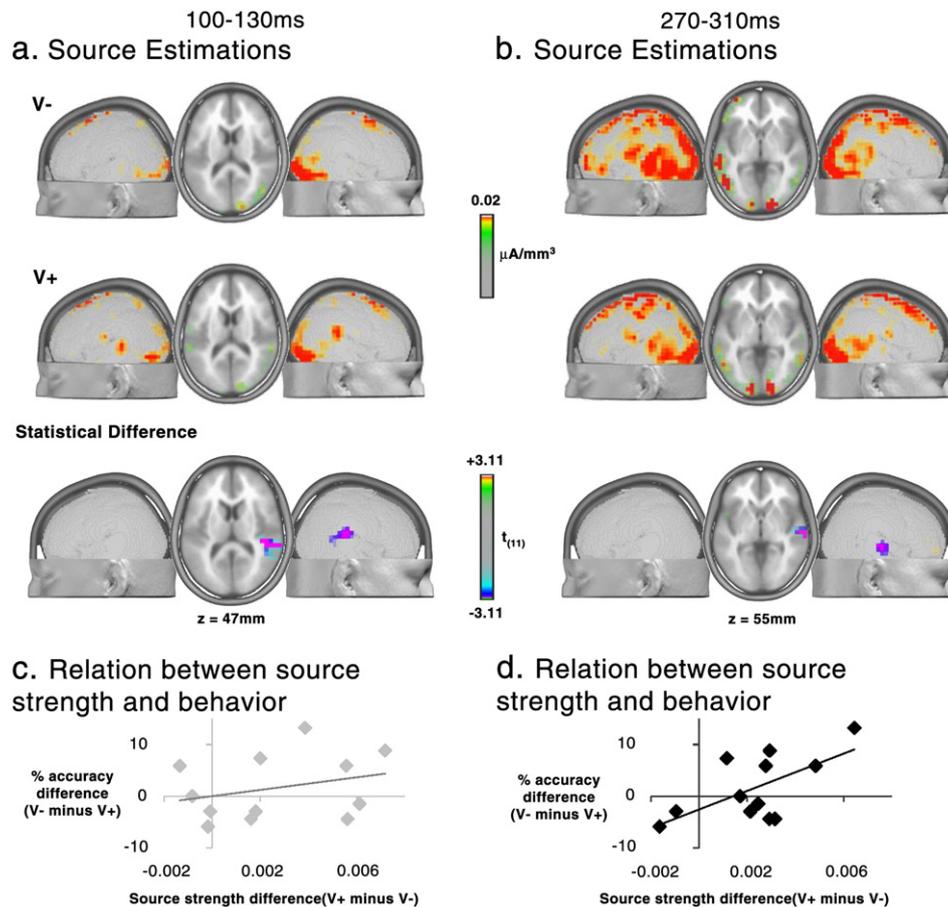
responses of both conditions, during both the 100–130 ms and 270–310 ms post-stimulus time periods two distinct maps were identified that appeared to differentially account for each condition (see Fig. 2e). This pattern observed in the group-averaged VEPs was statistically assessed in the single-subject VEPs using a spatial-correlation fitting procedure. There was a significant condition  $\times$  map interaction for both the 100–130 ms ( $F_{(1,11)} = 16.643$ ;  $p = 0.002$ ;  $\eta_p^2 = 0.602$ ) and the 270–310 ms ( $F_{(1,11)} = 6.348$ ;  $p = 0.028$ ;  $\eta_p^2 = 0.366$ ) post-stimulus periods. In both time periods one map predominated the responses to one of the conditions while the other predominated the responses to the other condition (see Fig. 2f).

The results to this point indicate that visual objects are incidentally discriminated according to the context of their initial presentation (i.e. in a unisensory or a meaningless multisensory context). Further they suggest that this differential discrimination takes place during early stages ( $\sim 100$  ms) of visual processing and engages different generator configurations.

#### Source estimations

Source estimations from both time periods were statistically analyzed to identify the likely brain regions contributing to these differential effects. During the early time period (100–130 ms post-stimulus onset) both conditions included prominent sources within the occipital,

temporo-parietal, and frontal lobes (upper portion of Fig. 3a). The statistical contrast of these source estimations identified a single cluster of solution points meeting our criteria (lower portion of Fig. 3a; see Materials and methods for details of statistical criteria). This cluster was located within the right posterior insular cortex (BA 13; coordinates of maximal  $t$ -value = 41,  $-28$ , 14 mm) and extended into the superior temporal gyrus (STG; BA 21). Although it did not meet our spatial extent criterion, there was evidence of a small 6-node cluster (BA 18; coordinates of maximal  $t$ -value 17,  $-87$ ,  $-9$  mm) within the right inferior lateral occipital cortex that exhibited stronger activity in response to the V $-$  than V $+$  condition. We mention this cluster here, given the previous evidence of differential activity within this region during this type of task (Murray et al., 2004, 2005). During the 270–310 ms post-stimulus period, both conditions again included prominent sources within the occipital, occipital-temporal and frontal lobes (upper portion of Fig. 3b). The statistical contrast again identified a single cluster meeting our significance criteria (lower portion of Fig. 3b). This cluster was found in the right STG (BA 22; coordinates of maximal  $t$ -value = 65,  $-17$ , 2 mm) and extended anteriorly to the middle temporal gyrus (MTG; BA 21). Source estimations in each of these clusters were then correlated with performance accuracy (Fig. 3c and d). There was no evidence for a significant correlation between source estimations over the 100–130 ms period and performance accuracy ( $r = 0.287$ ;  $t_{(10)} = 0.95$ ,  $p = 0.366$ ). By contrast, we observed a



**Fig. 3.** The results of the source estimations are shown. The columns show the results from the 100–130 ms (a) and the 270–310 ms (b) post-stimulus onset time windows, respectively. For each time period the mean activation for each condition (V $-$  and V $+$ ) are shown. Beneath, the  $t$ -value map of the contrast between conditions is shown. Note that the negative values (cooler colors) indicate that the clusters are more active in the V $+$  condition than in the V $-$  condition. Over the 100–130 ms post-stimulus period the maximal  $t$ -value was located at 41,  $-28$ , 14 mm. Over the 270–310 ms post-stimulus period the maximal  $t$ -value was located at 65,  $-17$ , 2 mm. (c and d) Correlations between the observed cost in discrimination accuracy ( $y$ -axis, in [%]) in the V $+$  condition compared to the V $-$  condition and the gain in source strength (predictor,  $x$ -axis, in [ $\mu\text{A}/\text{mm}^3$ ]). Only the later period (270–310 ms) yielded a significant correlation ( $r_{(10)} = 0.627$ ;  $p = 0.029$ ) between source strength in the in the mSTG/MTG cluster and the cost in discrimination accuracy over subjects.

significant positive correlation between the V+ vs. V– difference in source strength within the mSTG/MTG cluster and the cost in discrimination accuracy (V– minus V+) ( $r=0.627$ ;  $t_{(10)}=2.55$ ,  $p=0.029$ ).

## Discussion

The present study demonstrates that the discrimination of image repetitions is affected by whether or not the initial experience was multisensory versus solely visual. These effects occur even though the auditory information was both incidental for the task at hand and also devoid of any semantic content (i.e. is purely episodic in nature). This extends prior research showing this type of phenomenon following semantically congruent single-trial exposures to multisensory stimuli (Lehmann and Murray, 2005; Murray et al., 2004, 2005). We show that memory performance with visual objects is impaired by prior episodic multisensory pairings and cannot be explained by carry-over effects from the initial memory encoding or by an affect of differential attention (for a review see Murray and Sperdin, 2010). Electrical neuroimaging analyses revealed differential processing of repeated visual stimuli starting at ~100 ms post-stimulus onset and following from changes in the topography of the electric field at the scalp. Source estimations localized these effects to regions of the auditory cortex (and more subtly within the inferior LOC). Subsequent effects (270–310 ms) were again driven by topographic differences and we localized within auditory cortices of the STG/MTG. These later effects in turn positively correlated with behavior; individuals with larger differential STG/MTG responses exhibited larger costs in performance accuracy in indicating image repetitions. In what follows, we discuss our findings in terms of incidental effects of multisensory interactions on memory processes.

Our results show that single-trial multisensory learning occurs with the pairing of visual objects with meaningless sounds. Two factors contributing to this conclusion can be distinguished. First, semantic congruence between the senses is not a determinant factor. This conclusion was tentatively drawn by Lehmann and Murray (2005) who used a paradigm identical to that here, with the exception that the same sound (pure tone) was used for all multisensory pairings. This paradigm did not allow the authors to discern whether the effects derived from the episodic nature of the combination or instead from a potential “noising” of any single object representation (and its re-activation with image repetition) by the association of the same sound with multiple objects across the course of the experiment. The pairing of unique meaningless sounds in the present study allowed us to rule out such a “noising” mechanism and instead supports the role of episodic pairing in engendering distinct multisensory representations that can be incidentally accessed upon image repetition (something to which we return in our discussion of the electrical neuroimaging results). In this way, our findings somewhat challenge the hypothesis that single-trial multisensory interactions impact subsequent unisensory retrieval only in specific semantically congruent situations or when information across the senses is concordant. For example, some emphasize the role of either object familiarity (van der Linden et al., 2010) or ethological validity (von Kriegstein and Giraud, 2006) as principal factors in whether performance (and in some instances brain responses) would be affected by prior multisensory exposure.

Effects were reliably obtained following single-trial exposure and despite the task-irrelevance of the auditory stimulus. In this regard, effects of multisensory learning appear to occur in an incidental fashion. This conclusion originally came out from prior works using this paradigm (Lehmann and Murray, 2005; Murray et al., 2004, 2005). Similarly, Nyberg et al. (2000) showed that effects on auditory activity in response to repeated visual stimuli were not affected by the strength of the association between the senses (quantified by the number of repetitions of a given multisensory pairing). Another example of the effects of single-trial multisensory exposure on

subsequent unisensory processing is the recalibration of auditory spatial representations by vision (Wozny and Shams, 2011). More generally, these findings contribute to a growing literature emphasizing the both short- and long- lasting effects of multisensory interactions on subsequent unisensory processing (Meylan and Murray, 2007; Naue et al., 2011; Shams et al., 2011). One important aspect that will require additional data is the duration over which single-trial multisensory learning persists in its effects on later visual processing. Resolving this will undoubtedly impact the clinical/developmental applicability of our findings. It will also be beneficial for future research to clarify the determinants of whether multisensory pairings result in memory performance enhancements or decrements. On the one hand, our collective findings with a continuous recognition task suggest that semantic pairings reliably enhance memory performance whereas episodic pairings result in decrements. By contrast, explicit memory tasks are less consistent, such that performance decrements were observed with semantically congruent pairings (e.g. Nyberg et al., 2000). Clarifying the source(s) of this discrepancy will be necessary to derive potential utility of the single-trial and incidental nature of these effects in clinical and developmental populations; something at the focus of ongoing research within our group. It will likewise be important to determine the potential contribution of active learning of the multisensory associations. Prior studies using active learning sessions with similar paradigms have yielded mixed effects on performance, with some observing enhancements (von Kriegstein and Giraud, 2006) and others decrements (Butler and James, 2011; Nyberg et al., 2000; Wheeler and Petersen, 2000) for multisensory vs. unisensory (or within-modal) pairings.

The results also support there being a general time window wherein incoming visual stimuli are first incidentally discriminated according to past multisensory experiences. In the present as well as our prior study (Murray et al., 2004) VEPs to the V+ and V– conditions first differed at ~100 ms post-stimulus onset. Moreover, and in both studies, this effect was due to modulations in the VEP topography that in turn must follow from changes in the underlying generator configuration. This timing suggests that relatively early stages of visual object processing are subject to influences from past (single-trial) multisensory exposure and raises the question of the precise visual processes underway at this latency. Because the critical contrast in this study was between two sub-types of image repetitions that themselves required the same task-related analysis and motor response (i.e. both had to be recognized as “old”), any difference would presumably reflect processes subsequent to and/or complementing an initial (and potentially coarse-level) object recognition stage. Rapid visual object discrimination has indeed been reported, with effects as early as 100 ms (e.g. Liu et al., 2009; Rousselet et al., 2002; Thorpe et al., 1996; VanRullen and Thorpe, 2001), and effects of image repetition have been documented as early as ~50 ms post-stimulus onset (e.g. Michel et al., 2004; Mouchetant-Rostaing et al., 2000; Murray et al., 2009; Seeck et al., 1997). The present results extend these findings to show that discrimination of repeated objects can be impacted by whether or not their prior exposure occurred in a multisensory vs. unisensory context that was also completely task-irrelevant (and in fact likely ignored by subjects as suggested by their slower reaction times to the AV than V condition). Moreover, we show that this context can be purely episodic and need not be linked to the processing of semantic congruence or in fact any semantic information contained within the auditory channel (see also Butler and James, 2011). It is also worth noting that all images from both the V+ and V– condition were highly familiar, suggesting that object familiarity is not driving the present effects (van der Linden et al., 2010). In addition to the effect over the 100–130 ms post-stimulus period, we also observed significant topographic VEP modulations over the 270–310 ms post-stimulus period the sources of which significantly correlated with performance metrics.

The electrical neuroimaging analyses applied in this study also identified the likely sources contributing to the incidental discrimination of image repetitions as well as the link between brain activity and performance accuracy. In terms of source estimations, significant differences were observed within the right auditory cortices (STG) and to a lesser degree with object-related cortices of the right LOC during the 100–130 ms post-stimulus period. Source activity was significantly stronger for the V+ than V− condition within the STG and significantly weaker for the V+ than V− condition within the LOC. Differential responses within the LOC in a similar paradigm involving semantically congruent multisensory pairings were observed by Murray et al. (2004, 2005). Interestingly, in these studies responses were consistently stronger for the V+ than V− condition, which is the opposite directionality of what was observed in the present study. One way of reconciling these differences is that stronger activity may be associated with the condition resulting in more accurate discrimination performance. Such being said, we would note that there was no evidence of a reliable correlation between differences in performance and differences in source strength within the LOC ( $r_{(10)} = -0.287$ ;  $p > 0.35$ ). Over the 270–310 ms post-stimulus period, significantly stronger activity in response to the V+ condition was observed within auditory cortices (MTG). Effects within the auditory cortices (STG at 100–130 ms and MTG and 270–310 ms) were not reliably observed in our prior investigations.

The present observation of effects within nominally auditory regions in response to visual stimuli may be linked to the episodic nature of the multisensory pairings and/or the impaired performance for the V+ vs. V− condition. Support for the former can be found in fMRI studies showing that activity within superior temporal regions is inversely related to the strength of the association between arbitrary auditory-visual multisensory stimulus combinations (e.g. Tanabe et al., 2005; see also Naghavi et al., 2011 for effects during encoding that are in turn linked with subsequent memory performance). Support for the latter possibility can be gleaned from the results of Nyberg et al. (2000). These authors observed stronger responses within auditory cortices in response to visually presented words that had been explicitly learned and remembered as being previously paired with their corresponding sounds. Interestingly and somewhat downplayed by Nyberg et al., there was an 8% performance decrement for words previously paired with sounds vs. unpaired words that was not statistically assessed. It may therefore be the case that the enhanced auditory activity reflects this performance cost and by extension failed retrieval processes. Another possibility is that these enhanced auditory cortex responses reflect memory processes linked to the reactivation of the initial context (i.e. redintegration; Hamilton, 1859; see also ; Wheeler and Petersen, 2000). As we found no evidence of a correlation between performance decrements and differential source activity within the STG over the 100–130 ms post-stimulus period, we can neither support nor refute any causal link between differential activity within the STG and performance decrements. By contrast, there was a significant correlation between performance decrements and differential source activity within the MTG over the 270–310 ms post-stimulus period. The more strongly this cluster was active for the V+ than the V− condition, the larger the performance difference was between the V− vs. V+ condition.

Despite this link between brain activity and performance, our use of episodic pairings makes an interpretation in terms of pure redintegration unlikely. Participants were engaged in a visually demanding task and showed no benefit from the auditory information, but rather performance impairments. Additionally, because the sounds were unique and meaningless, it is unlikely that participants were able to establish a distinct representation for each sound based on single-trial exposure. Instead, it may be the case that the auditory cortex is activated in an unspecific manner in response to images that had been paired with sounds; a form of echoic memory elicited by images. A more convincing demonstration of stimulus-specific redintegration

is found in von Kriegstein and Giraud (2006), who showed enhanced activity within the fusiform face area in response to voices that had been associated with faces but not voices that had been paired with names. More recently, it has been shown that responses to visual stimuli within auditory cortices differ according to the semantic category of the visual stimulus (e.g. animals, musical instruments, etc.; Meyer et al., 2010). Although the role of mental imagery cannot be fully discounted, these types of data nonetheless suggest that low-level cortices may respond in a semi-selective manner to stimuli from other sensory modalities and in particular to stimuli with multisensory associations. An alternative account of the present results is that episodic multisensory experiences of the kind used here elicit recursive activity within auditory cortices that is disadvantageous for the discrimination of image repetitions. In this way, the present findings may offer a potential compromise in that varieties of redintegration processes may elicit distinct networks as a function of performance accuracy. That is, situations leading to improved performance (e.g. following semantically congruent exposures as in Murray et al., 2004, 2005) would recruit a network predominantly within lateral occipital cortices (and perhaps also intraparietal sulcus; e.g. Werner and Noppeney, 2010). By contrast, situations leading to impaired performance (e.g. following meaningless and episodic exposures) would recruit a network predominantly within auditory cortices (both STG and MTG). To more fully assess this possibility will require contrasting V+ and V− conditions as a function of performance accuracy; something that would require sufficient numbers of trials leading to inaccurate memory discrimination. This was unfortunately not the case in the present study.

Multiple temporal phases and/or levels of differential activity have been previously observed in studies of multisensory object discrimination (e.g. Diaconescu et al., 2011; Kayser, 2010; Molholm et al., 2004; Schneider et al., 2008; Werner and Noppeney, 2010) and multisensory interactions between simple stimuli (e.g. Cappe et al., 2010, 2012). In one model, Werner and Noppeney (2010) proposed that responses within primary auditory cortices are enhanced by (corresponding) visual stimuli independently of task-context and without a direct link to performance metrics. These effects were considered as a general mechanism of multisensory enhancement of stimulus salience without a direct relationship with behavior. Effects within the superior temporal sulcus, planum temporale, and inferior parietal sulcus were significantly correlated with performance gains on object classification and were thus considered by Werner and Noppeney (2010) to be involved in the integration of object-specific features. The extent to which our results can be grafted onto this model is not immediately forthcoming and will undoubtedly require further experimental data. This is particularly the case because responses within auditory cortices were not observed in our prior studies involving semantically congruent multisensory learning (Murray et al., 2004, 2005); something that the Werner and Noppeney (2010) model might otherwise have predicted. Likewise, the above model is based on interactions between externally presented and semantically congruent objects stimuli, whereas the present results reflect the influence of past multisensory experiences on current unisensory visual processing. Nonetheless, it is interesting to note the parallel between the effects obtained by Werner and Noppeney (2010) and our observations of behaviorally independent responses in relatively low-level auditory cortices at 100–130 ms followed by behaviorally coupled responses in higher-order auditory regions at 270–310 ms post-stimulus onset. One possibility is that prior multisensory exposures (at least those that are episodic in nature) manifest themselves on current visual processing in a manner similar to what occurs following the presentation of actual multisensory stimuli. Assuming such, the responses within auditory cortices (both STG and MTG) may serve to facilitate the differentiation of incoming visual stimuli even if ultimately to the detriment of memory performance accuracy.

Supplementary materials related to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2012.05.027>.

## Acknowledgments

This work has been supported by the Swiss National Science Foundation (grant 310030B-133136 to MMM). The Cartool software has been programmed by Denis Brunet, from the Functional Brain Mapping Laboratory, Geneva, Switzerland, and is supported by the EEG Brain Mapping Core of the Center for Biomedical Imaging ([www.cibm.ch](http://www.cibm.ch)) of Geneva and Lausanne. It is freely available at: <http://sites.google.com/site/fbmlab/cartool/cartooldownload>. Jean-François Knebel provided input on the statistical analysis on spectra and luminance of our visual stimuli. Marzia De Lucia assisted with additional computer programming. Rosanna De Meo provided helpful comments on a preliminary version of this manuscript.

## References

- Baier, B., Kleinschmidt, A., Müller, N.G., 2006. Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *J. Neurosci.* 26 (47), 12260–12265.
- Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011, 813–870.
- Butler, A.J., James, K.H., 2011. Cross-modal versus within-modal recall: differences in behavioral and brain responses. *Behav. Brain Res.* 224, 387–396.
- Calvert, G.A., 2004. *The handbook of multisensory processes*. MIT Press, Cambridge, MA.
- Cappe, C., Thut, G., Romei, V., Murray, M.M., 2010. Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *J. Neurosci.* 30 (38), 12572–12580.
- Cappe, C., Thelen, A., Romei, V., Thut, G., Murray, M.M., 2012. Looming signals reveal synergistic principles of multisensory integration. *J. Neurosci.* 32 (4), 1171–1182.
- De Lucia, M., Clarke, S., Murray, M.M., 2010a. A temporal hierarchy for conspecific vocalization discrimination in humans. *J. Neurosci.* 30 (33), 11210–11221.
- De Lucia, M., Michel, C.M., Murray, M.M., 2010b. Comparing ICA-based and single-trial topographic ERP analyses. *Brain Topogr.* 23 (2), 119–127.
- Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topogr.* 14 (2), 131–137.
- Grave de Peralta Menendez, R., Murray, M.M., Michel, C.M., Martuzzi, R., Gonzalez Andino, S.L., 2004. Electrical neuroimaging based on biophysical constraints. *Neuroimage* 21 (2), 527–539.
- Diaconescu, A.O., Alain, C., McIntosh, A.R., 2011. The co-occurrence of multisensory facilitation and cross-modal conflict in the human brain. *J. Neurophysiol.* 106, 2896–2909.
- Gonzalez Andino, S.L., Murray, M.M., Foxe, J.J., de Peralta Menendez, R.G., 2005. How single-trial electrical neuroimaging contributes to multisensory research. *Exp. Brain Res.* 166 (3–4), 298–304.
- Gottfried, J.A., Smith, A.P., Rugg, M.D., Dolan, R.J., 2004. Remembrance of odors past: human olfactory cortex in cross-modal recognition memory. *Neuron* 42 (4), 687–695.
- Guthrie, D., Buchwald, J.S., 1991. Significance testing of difference potentials. *Psychophysiology* 28 (2), 240–244.
- Hamilton, W., 1859. *Lectures on Metaphysics and Logic*. Gould & Lincoln, Boston.
- Kayser, C., Logothetis, N.K., Panzeri, S., 2010. Visual enhancement of the information representation in auditory cortex. *Curr. Bio.* 20 (1), 19–24.
- Kim, R.S., Seitz, A.R., Shams, L., 2008. Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One* 3 (1), e1532.
- Knebel, J.F., Murray, M.M., 2012. Towards a resolution of conflicting models of illusory contour processing in humans. *Neuroimage* 59 (3), 2808–2817.
- Knebel, J.F., Toepel, U., Hudry, J., le Coutre, J., Murray, M.M., 2008. Generating controlled image sets in cognitive neuroscience research. *Brain Topogr.* 20 (4), 284–289.
- Knebel, J.F., Javitt, D.C., Murray, M.M., 2011. Impaired early visual response modulations to spatial information in chronic schizophrenia. *Psychiatry Res.* 193 (3), 168–176.
- Koenig, T., Melie-Garcia, L., 2010. A method to determine the presence of averaged event-related fields using randomization tests. *Brain Topogr.* 23 (3), 233–242.
- Lehmann, D., 1987. Principles of spatial analysis. In: Gevins, A.S., Remond, A. (Eds.), *Handbook of electroencephalography and Clinical Neurophysiology*. : Methods of Analysis of Brain Electrical and Magnetic Signals, vol. 1. Elsevier, Amsterdam, pp. 309–354.
- Lehmann, S., Murray, M.M., 2005. The role of multisensory memories in unisensory object discrimination. *Brain Res. Cogn. Brain Res.* 24 (2), 326–334.
- Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48, 609–621.
- Liu, H., Agam, Y., Madsen, J.R., Kreiman, G., 2009. Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62 (2), 281–290.
- Martuzzi, R., Murray, M.M., Meuli, R.A., Thiran, J.P., Maeder, P.P., Michel, C.M., Grave de Peralta Menendez, R., Gonzalez Andino, S.L., 2009. Methods for determining frequency- and region-dependent relationships between estimated LFPs and BOLD responses in humans. *J. Neurophysiol.* 101 (1), 491–502.
- Meyer, K., Kaplan, J.T., Essex, R., Webber, C., Damasio, H., Damasio, A., 2010. Predicting visual stimuli on the basis of activity in auditory cortices. *Nat. Neurosci.* 13 (6), 667–668.
- Meylan, R.V., Murray, M.M., 2007. Auditory-visual multisensory interactions attenuate subsequent visual responses in humans. *Neuroimage* 35 (1), 244–254.
- Michel, C.M., Murray, M.M., 2012. Towards the utilization of EEG as a brain imaging tool. *Neuroimage*, <http://dx.doi.org/10.1016/j.neuroimage.2011.12.039>.
- Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. *Clin. Neurophysiol.* 115 (10), 2195–2222.
- Molholm, S., Ritter, W., Javitt, D.C., Foxe, J.J., 2004. Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cereb. Cortex* 14, 452–465.
- Mouchetant-Rostaing, Y., Giard, M.H., Delpuech, C., Echallier, J.F., Pernier, J., 2000. Early signs of visual categorization for biological and non-biological stimuli in humans. *Neuroreport* 11 (11), 2521–2525.
- Murray, M.M., Sperdin, H.F., 2010. Single-trial multisensory learning and memory retrieval. In: Kaiser, J., Naumer, M.J. (Eds.), *Multisensory object perception in the primate brain*, <http://dx.doi.org/10.1007/978-1-4419-5615-11>.
- Murray, M.M., Michel, C.M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., Schneider, A., 2004. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* 21 (1), 125–135.
- Murray, M.M., Foxe, J.J., Wylie, G.R., 2005. The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage* 27 (2), 473–478.
- Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20 (4), 249–264.
- Murray, M.M., De Santis, L., Thut, G., Wylie, G.R., 2009. The costs of crossing paths and switching tasks between audition and vision. *Brain Cogn.* 69 (1), 47–55.
- Naghavi, H.R., Eriksson, J., Larsson, A., Nyberg, L., 2011. Cortical regions underlying successful encoding of semantically congruent and incongruent associations between common auditory and visual objects. *Neurosci. Lett.* 505 (2), 191–195.
- Naue, N., Rach, S., Struber, D., Huster, R.J., Zaehle, T., Korner, U., Herrmann, C.S., 2011. Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *J. Neurosci.* 31 (21), 7729–7736.
- Nyberg, L., Habib, R., McIntosh, A.R., Tulving, E., 2000. Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl. Acad. Sci. U. S. A.* 97 (20), 11120–11124.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1), 97–113.
- Perrin, F., Pernier, J., Bertrand, O., Giard, M.H., Echallier, J.F., 1987. Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr. Clin. Neurophysiol.* 66 (1), 75–81.
- Pourtois, G., Delplanque, S., Michel, C., Vuilleumier, P., 2008. Beyond conventional event-related brain potential (ERP): exploring the time-course of visual emotion processing using topographic and principal component analyses. *Brain Topogr.* 20 (4), 265–277.
- Raij, T., Ahveninen, J., Lin, F.H., Witzel, T., Jääskeläinen, I.P., Letham, B., Israeli, E., Sahyoun, C., Vasilos, C., Stufflebeam, S., Hämäläinen, M., Belliveau, J.W., 2010. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur. J. Neurosci.* 31 (10), 1772–1782.
- Ranganath, C., Rainer, G., 2003. Neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci.* 4 (3), 193–202.
- Rousselle, G.A., Fabre-Thorpe, M., Thorpe, S.J., 2002. Parallel processing in high-level categorization of natural images. *Nat. Neurosci.* 5 (7), 629–630.
- Rugg, M.D., Johnson, J.D., Park, H., Uncapher, M.R., 2008. Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. *Prog. Brain Res.* 169, 339–352.
- Schneider, T.R., Debener, S., Oostenveld, R., Engel, A.K., 2008. Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *Neuroimage* 42 (3), 1244–1254.
- Seeck, M., Michel, C.M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., Landis, T., Schomer, D.L., 1997. Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport* 8 (12), 2749–2754.
- Shams, L., Seitz, A.R., 2008. Benefits of multisensory learning. *Trends Cogn. Sci.* 12 (11), 411–417.
- Shams, L., Wozny, D.R., Kim, R., Seitz, A., 2011. Influences of multisensory experience on subsequent unisensory processing. *Front. Psychol.* 2 (264).
- Snodgrass, J.G., Vanderwart, M., 1980. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6 (2), 174–215.
- Spence, C., Nicholls, M.E., Driver, J., 2001. The cost of expecting events in the wrong sensory modality. *Percept. Psychophys.* 63 (2), 330–336.
- Stein, B.E., Meredith, M.A., 1993. *The merging of the senses*. The MIT Press, Cambridge Mass.; London.
- Talairach, J., Tournoux, P., 1988. *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. G. Thieme Medical Publishers, Stuttgart; New York.
- Tanabe, H.C., Honda, M., Sadato, N., 2005. Functionally segregated neural substrates for arbitrary audiovisual paired-association learning. *J. Neurosci.* 25 (27), 6409–6418.
- Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. *Nature* 381 (6582), 520–522.

- Toepel, U., Knebel, J.F., Hudry, J., le Coutre, J., Murray, M.M., 2009. The brain tracks the energetic value in food images. *Neuroimage* 44 (3), 967–974.
- Tzovara, A., Murray, M.M., Michel, C.M., De Lucia, M., in press. A tutorial review of electrical neuroimaging from group-average to single-trial event-related potentials. *Dev. Neuropsychol.*
- van der Linden, M., van Turenout, M., Indefrey, P., 2010. Formation of category representations in superior temporal sulcus. *J. Cogn. Neurosci.* 22 (6), 1270–1282.
- VanRullen, R., Thorpe, S.J., 2001. Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception* 30 (6), 655–668.
- von Kriegstein, K., Giraud, A.L., 2006. Implicit multisensory associations influence voice recognition. *PLoS Biol.* 4 (10), e326.
- Wheeler, M.E., Petersen, S.E., 2000. Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc. Natl. Acad. Sci. U. S. A.* 97 (20), 11125–11129.
- Werner, S., Noppeney, U., 2010. Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *J. Neurosci.* 30 (7), 2662–2675.
- Wozny, D.R., Shams, L., 2011. Computational characterization of visually induced auditory spatial adaptation. *Front. Integr. Neurosci.* 5 (75).