

## A multisensory perspective on object memory

Pawel J. Matusz<sup>a</sup>, Mark T. Wallace<sup>b, c, d, e, f</sup>, Micah M. Murray<sup>a, b, g, h, \*</sup>

<sup>a</sup> The Laboratory for Investigative Neurophysiology (The LINE), Neuropsychology & Neurorehabilitation Service & Department of Radiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland

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

<sup>c</sup> Department of Psychology, Vanderbilt University, Nashville, TN, USA

<sup>d</sup> Kennedy Center for Research on Human Development, Vanderbilt University, Nashville, TN, USA

<sup>e</sup> Department of Psychiatry, Vanderbilt University, Nashville, TN, USA

<sup>f</sup> Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA

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

<sup>h</sup> Department of Ophthalmology, University of Lausanne, Jules-Gonin Eye Hospital, Lausanne, Switzerland

### ARTICLE INFO

#### Keywords:

Multisensory  
Cross-modal  
Auditory  
Visual  
Object  
Memory  
Learning

### ABSTRACT

Traditional studies of memory and object recognition involved objects presented within a single sensory modality (i.e., purely visual or purely auditory objects). However, in naturalistic settings, objects are often evaluated and processed in a multisensory manner. This begets the question of how object representations that combine information from the different senses are created and utilised by memory functions. Here we review research that has demonstrated that a single multisensory exposure can influence memory for both visual and auditory objects. In an old/new object discrimination task, objects that were presented initially with a task-irrelevant stimulus in another sense were better remembered compared to stimuli presented alone, most notably when the two stimuli were semantically congruent. The brain discriminates between these two types of object representations within the first 100 ms post-stimulus onset, indicating early “tagging” of objects/events by the brain based on the nature of their initial presentation context. Interestingly, the specific brain networks supporting the improved object recognition vary based on a variety of factors, including the effectiveness of the initial multisensory presentation and the sense that is task-relevant. We specify the requisite conditions for multisensory contexts to improve object discrimination following single exposures, and the individual differences that exist with respect to these improvements. Our results shed light onto how memory operates on the multisensory nature of object representations as well as how the brain stores and retrieves memories of objects.

### 1. Introduction

Imagine that you are at a cocktail party and you are being introduced by a friend to a group of strangers. Let's call them Sarah, Kim and Deborah. Your friend introduces you and tells the group that you are a cognitive neuroscientist who is visiting town. During the next two minutes, you exchange a few sentences with Sarah. During the same two minutes, you will only see Kim smiling politely when shaking your hand, and you will not happen to hear Deborah introducing herself to you, as someone behind her will shout loudly to his friend standing in the other corner of the room (Fig. 1).<sup>1</sup> A week after this cocktail party

you are at a different gathering, where you once again see Sarah, Kim, and Deborah. Whose face you will recognise more easily?

Psychophysical, neurophysiological, and human brain imaging research over the last 40 years has greatly advanced our understanding of the cognitive and brain mechanisms that support perception and memory as well as the interactions that they share in everyday situations (Constantinescu et al., 2016; Gazzaley and Nobre, 2012; Mahon and Caramazza, 2011; Summerfield and de Lange, 2014). In such everyday situations, when we encounter a new person or a new object, information about them is typically conveyed by more than a single sense. Indeed, under such multisensory circumstances, profound

\* Corresponding author at: The Laboratory for Investigative Neurophysiology (The LINE), Neuropsychology & Neurorehabilitation Service & Department of Radiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland.

Email address: Micah.murray@chuv.ch (M.M. Murray)

changes in behaviour and perception can be elicited, and these changes are accompanied by striking changes in the patterns of brain activation and the networks that are engaged. Auditory-visual multisensory processes have been identified throughout functional cortical hierarchies, including primary cortices (reviewed in Murray et al. (2016a)) infero-temporal and superior temporal regions (reviewed in Lewis (2010) for the case of auditory-visual object processing) as well as prefrontal regions (reviewed in Murray and Wallace (2012)). Although much emphasis has been placed on behavioural and perceptual processes, recent work has also shown that the presentation of sensory stimuli in a multisensory manner can also have profound effects on our memories, and provide important clues as to why you can recognise Sarah better than her friends on your second meeting in the example provided above. In the current review, we discuss the evidence that multisensory context can improve unisensory object discrimination even after a single exposure. We then specify the requisite conditions for such improvements as well as the individual differences therein. Lastly, we place the reviewed behavioural and brain imaging findings within the broader literature on multisensory learning and discuss the importance of considering multisensory contributions when creating accurate models of object perception and memory.

While the experimental paradigm that we have employed has been described in detail previously (Thelen and Murray, 2013; for a summary, see Fig. 2a), we summarise it here briefly. We employed a continuous recognition task, in which on each trial participants have to indicate as quickly and accurately as possible whether they saw a given object for the first (“new”) or second (“old”) time. Across different variations of this paradigm utilised in a number of studies over the years, stimuli within one sense (e.g. vision) would always be task-relevant, while stimuli in another sense (e.g. audition) would always be task-irrelevant. The initial and repeated trials were always equally probable, and across all trial types the number of unisensory and multisensory trials were also equally probably distributed. While it was the case that some of our early work involved a paradigm where multisensory information was only presented on initial trials, subsequent work has replicated effects even when rendering the multisensory content uninformative about the task-relevant dimension (i.e. whether an object was presented for the initial or repeated time). In this line of research, the effectiveness of three distinct multisensory contexts in improving memory has been assessed: 1) a semantically congruent context – where the task-relevant and task-irrelevant stimuli represent the same object (e.g., a drawing of a cow combined with a sound “moo”), 2) a meaningless-association context – where the task-relevant stimuli are paired with tones or noises, and 3) a semantically incongruent context – where the task-relevant and task-irrelevant stimuli represent different objects.

By manipulating the number (and type) of senses actively engaged, the nature of the relationship between the stimuli across the two senses, as well as their task-relevance, our paradigm sought to more closely emulate information processing in naturalistic environments. This evidence (and more recently that from other independent laboratories) has provided novel insights into the behavioural and brain mechanisms guiding memory and information processing in everyday situations. The overall message from these studies is that memory for objects is generally improved when the information is first encountered in a multisensory manner.

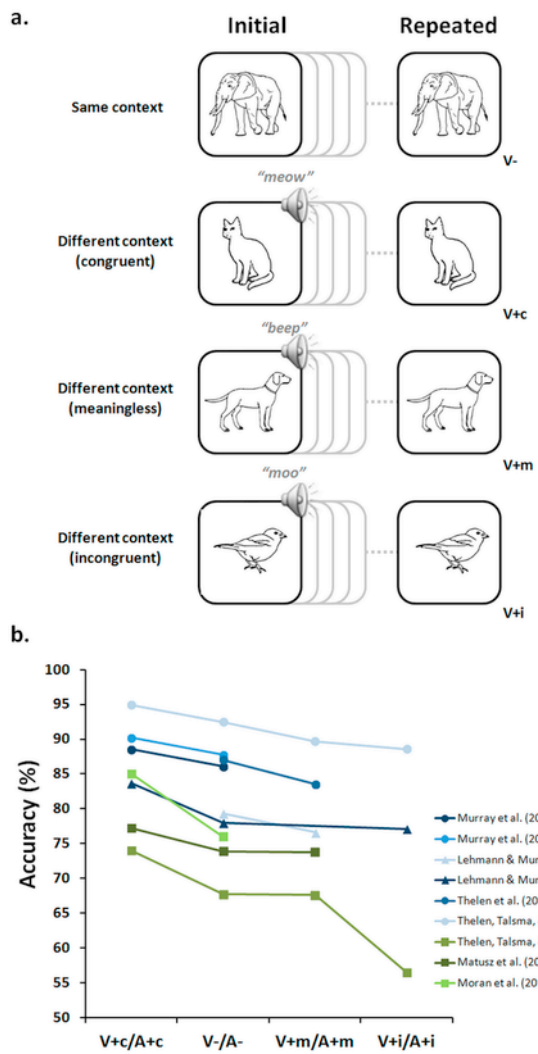


Fig. 1. A cartoon of a cocktail party setting. This is a typical scenario where multisensory information that is synchronous, co-localised and semantically-congruent co-occurs with information that is none of these. It is also exemplary of a scenario where information must be learned for later recognition in a different context.

## 2. Which multisensory contexts improve memory?

In our paradigm, the benefits on object memory of having information presented in a multisensory manner are generally observed as improved discrimination accuracy. Reaction times (RTs) showed no similar benefits (e.g., Lehmann and Murray, 2005). When the initial multisensory presentation (and encoding) involved semantically congruent pairings, robust memory improvements were observed on subsequent retrieval. These improvements were observed across studies employing different stimulus and paradigm parameters that balanced the occurrence of multisensory information on initial and repeated presentations, and distinct brain mapping methods. For example, these improvements in discrimination were seen in Murray et al.’s (2005) fMRI study, despite the presence of scanner noise that arguably could have interfered with the ability to perceive/encode the task-irrelevant sounds and also extended the usual item repetition lag of 5 s (used in the psychophysical and EEG studies) up to 50 s to accommodate inter-trial intervals necessary due to the constraints of fMRI data acquisition (see Table 1). Across studies, the observed benefits for semantically congruent multisensory pairings on memory performance ranged from a gain of 2.5–9% over performance on unisensory visual or auditory trials (also Fig. 2b). Where the study design enabled the calculation of a more rigorous measure of sensory processing (the perceptual sensitivity parameter,  $d'$ , Macmillan and Creelman, 2004), these multisensory benefits were found to be even larger (i.e., 12% performance memory improvement; Matusz et al., 2015b). Overall, these improvements have been seen across 6 studies involving more than 100 participants and exhibiting effect sizes ranging from small to large ( $\eta^2_p = 0.14-0.63$ ; see Table 1 of Thelen and Murray, 2013 for details; see

<sup>1</sup> Please note that while this scenario may provide a good approximation of the multisensory effects on memory, the more social and attention-demanding nature of person-to-person interactions render it somewhat different from the paradigm/s we have focused on in this review.



**Fig. 2.** a. Schematic of the multisensory continuous recognition task. When vision is the task-relevant sensory modality, the participant indicates if the image is being presented for the first or a repeated time. Initial presentations are divided between those that are unisensory visual and those which are multisensory. The multisensory context varies according to the semantic content of the sound (here congruent, meaningless, or incongruent). Repeated presentations are exclusively visual and therefore differ only in how they had been initially experienced (denoted by V-, V+c, V+m, and V+i). In a block of trials, all of these stimulus conditions are inter-mixed. b. Summary of behavioural findings. Accuracy for the various repeated presentations are displayed. The blue lines refer to studies where the task was performed in the visual modality, while green lines refer to studies where the task was performed in the auditory modality. Across studies, it can be seen that stimuli that had been initially presented in a semantically congruent multisensory context result in higher accuracy than stimuli that had only been experienced in a unisensory context. Other had-been multisensory contexts generally result in no difference or even performance impairment relative to the unisensory context.

also Moran et al. (2013); for similar size of effects in studies involving setups with separate exposure and recall, see Heikkilä et al. (2015), Heikkilä and Tiippana (2016), Naghavi et al. (2011), Ueno et al. (2015)).

In contrast to when multisensory stimuli were semantically congruent, if the initial pairing is semantically incongruent, the typical result is memory *impairments* relative to when stimuli are initially presented in a unisensory manner, with the impairments ranging from between a 4% and 16.5% decrease in discrimination accuracy (Fig. 2b). Similarly, if the initial presentations involved pairings with meaningless task-irrelevant information, performance decrements of 3–4% were typically seen (Fig. 2b). Intriguingly, for these meaningless pairings, perfor-

**Table 1**  
Factors influencing memory improvements based on multisensory processes.

Factor	Description	Sufficiency for memory improvement?	Support
Implicitness	The multisensory nature of the stimuli is task-irrelevant.	✓	Murray et al. (2004, 2005); Gottfried et al. (2004); Lehmann and Murray (2005); Naghavi et al. (2011); Moran et al. (2013); Thelen et al. (2012, 2015); Thelen et al. (2014); Matusz et al. (2015a); Heikkilä et al. (2015, 2016)
Continuous recognition	Study and test items are intermixed.	✓	Murray et al. (2004, 2005); Lehmann and Murray (2005); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen et al. (2014); Matusz et al. (2015a)
Single-trial	Initial multisensory presentations occur once.	✓	Murray et al. (2004, 2005); Lehmann and Murray (2005); Naghavi et al. (2011); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen et al. (2014); Matusz et al. (2015a)

Table 1 (Continued)

Factor	Description	Sufficiency for memory improvement?	Support
Semantic congruence	The initial presentation consists of stimuli referring to the same object.	✓	Murray et al. (2004, 2005); Lehmann and Murray (2005; Exp2); Naghavi et al. (2011); von Kriegstein et al. (2006); Moran et al. (2013); Thelen et al. (2015); Matusz et al. (2015a); Ueno et al. (2015); Heikkilä et al. (2015, 2016)
Object-ness	The stimuli involve naturalistic or schematized images/sounds or visual/spoken object labels	✓	Murray et al. (2004, 2005); Gottfried et al. (2004); Lehmann and Murray (2005); von Kriegstein et al. (2006); Naghavi et al. (2011); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen et al. (2014); Matusz et al. (2015a); Ueno et al. (2015); Heikkilä et al. (2015, 2016)

Table 1 (Continued)

Factor	Description	Sufficiency for memory improvement?	Support
Variation in multisensory pairings	Stimuli in the task-irrelevant sense vary across trials	? <sup>a</sup>	Murray et al. (2004, 2005); Gottfried et al. (2004); Lehmann and Murray (2005; Exp2); von Kriegstein et al. (2006); Naghavi et al. (2011); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen et al. (2014); Matusz et al. (2015a); Ueno et al. (2015); Heikkilä et al. (2015, 2016)
Duration/time	Time interval between initial and repeated item (in continuous recognition)	Within the same block/ up to 1 min	Murray et al. (2004, 2005); Lehmann and Murray (2005; Exp2); Thelen et al. (2015); Moran et al. (2013); Thelen et al. (2014); Matusz et al. (2015a)
Persistence over multiple intervening items	Benefits extend beyond a single intervening item (in continuous recognition)	Average 13±3 items	Murray et al. (2004, 2005); Lehmann and Murray (2005; Exp2); Thelen et al. (2015); Moran et al. (2013); Thelen et al. (2014); Matusz et al. (2015a);
Persistence over multiple sessions	Benefits extend beyond a single testing session (in continuous recognition)	?	–

Table 1 (Continued)

Factor	Description	Sufficiency for memory improvement?	Support
Physical vs. conceptual initial-repeated semantic congruence	Initial and repeated presentations are physically identical or refer to the same object, parametrically varying semantic relatedness	?	–

<sup>a</sup> This factor exhibits high inter-individual variability.

mance was highly variable across individuals. Thus, approximately half of the tested participants demonstrated memory improvements following these meaningless initial contexts, and these improvements were seen for both visual and auditory memory.<sup>2</sup> For visual memory, the improvements ranged between 0.5% and 7% (Thelen et al., 2014), while for auditory memory these gains were between 2.5% and 10.8% (Thelen et al. 2014, Supplemental Information). These findings suggest that multisensory pairings involving merely the simultaneity of stimulus onsets across the senses may only be weakly effective in supporting object memory (cf., De Meo et al., 2015; Murray et al., 2016a, 2016b; ten Oever et al., 2016 for reviews on the role of audiovisual simultaneity detection in modulating instantaneous perception and selective attention). We discuss these results in more detail below in the section on individual differences (Section 3).

To summarise, in a continuous discrimination paradigm, the initial presentation of stimuli in a multisensory context (as opposed to unisensory), whether congruent or incongruent, has significant influence on memory performance. Across studies, we have identified several requisite conditions under which initial multisensory presentation improve memory (Table 1). These conditions, and the effects that we observed as they were manipulated, challenge some of the basic tenets of traditional models of memory. First, one of the most established findings in this domain is that memory performance is best in situations where the encoding and retrieval contexts are identical (Baddeley et al., 2009; Smith and Vela, 2001). Our results strongly suggest that those findings hold true mainly, if not exclusively, in unisensory settings. Naturalistic environments, whether it be a classroom or a cocktail party, are typically multisensory in nature. According to our results, in such naturalistic environments, any memory benefits can be further enhanced in cases where the encoded object stimulated multiple senses. If the signal in the other sense semantically matches the identity of the object presented in the task-relevant sense and, thus, activate presumably long-term memory associations, these benefits will likely be visible in all individuals. We would emphasize that attention was always focused on one sense exclusively. Yet, the benefits of multisensory presentations were nonetheless observed and thus can be considered implicit. This highlights the efficacy of multisensory processes in influencing object memory even in situations where they occur outside of the focus of selective attention.

An important issue is to what extent these memory benefits are specific to multisensory presentations. To our knowledge, this has not yet been specifically investigated with a continuous recognition paradigm similar to that which we have used. That said, there are behavioural

<sup>2</sup> We would note that a similar degree of inter-individual variability is observed with semantically incongruent pairings, though to date this has not been specifically investigated with brain mapping/imaging methods.

data in cats showing that performance enhancements are greater for multisensory than for unisensory redundancy (Gingras et al., 2009). Likewise, multisensory benefits were found to be greater than unisensory benefits during a masked letter identification task in humans (Chen and Spence, 2011). Finally and most germane are data from two studies. One study had non-human primates performing a delayed match-to-sample task (Gibson and Maunsell, 1997). They showed that the propensity of selective delay period activity at IT neurons was significantly greater for multisensory than for unisensory learned associations. The other study had human participants perform a recognition memory task with separate learning and memory phases (Heikkilä et al., 2015). Semantic congruence at encoding was beneficial for later recognition memory when the materials were multisensory, but not when they were unisensory visual. Collectively these results, alongside our own data, would support an account based on multisensory processing.

Multisensory interactions based on temporal relations (i.e., onset simultaneity) or on well-learned associations can influence memory despite the multisensory aspect being task-irrelevant and thus outside the goals of the observer. The presence of these multisensory-based memory benefits has important clinical implications for treatment and rehabilitation of memory and sensory disorders (e.g. Johansson, 2012; Baum et al., 2015), and, thus, an important question is how generalisable these results are. Our ongoing work, which demonstrated a link between the strength of multisensory benefits on a simple reaction time task and a degree of preservation of higher-level functioning assessed with a standardised questionnaire in individuals with a mild cognitive impairment, is focusing specifically on this question (Eardley et al., in preparation). While we discuss the prerequisite conditions for these memory modulations to occur in Section 3, we will now discuss how the observed memory effects change as a function of the task-relevant sense.

The last several years have seen a growing interest in the role of the task-relevant sense as a bottom-up factor modulating multisensory processing (Romei et al., 2009, 2013; Schmid et al., 2011). Auditory object memory is generally reported to be weaker than visual memory (Cohen et al., 2009; Yuval-Greenberg and Deouell, 2007, 2009). This poses an important question of whether benefits from encoding stimuli in a multisensory context would be larger for auditory than visual memory. Such a finding would be similar to observations that stronger benefits of multisensory processing are frequently observed in situations in which the inputs are weakly effective (“inverse effectiveness principle”, Crosse et al., 2016; Jiang et al., 2001; Stevenson et al., 2014; Stevenson and James, 2009; Wallace, 2004; Wallace et al., 2006; Wallace et al., 1998; Wallace et al., 2004). Evidence in support of this notion was found in one of our studies (Thelen et al., 2015), where the same group of participants performed the old/new task first with visual and then with auditory objects as task-relevant objects, or vice versa. As in other studies (Cohen et al., 2009; Yuval-Greenberg and Deouell, 2007, 2009), auditory memory was generally much weaker than visual memory (67% vs. 92% accuracy, respectively). Similarly to our previous studies, we found memory benefits that were exclusive for semantically congruent pairings, and that were seen for both vision and hearing. Notably, the auditory memory benefits were approximately four times larger than the visual benefits (i.e., 8.8% vs. 2.2% improvement in object discrimination). Because a very similar setup was used across the two tasks within the same individuals, our findings strongly suggest that the rule by which pairings involving “less effective” inputs trigger stronger multisensory processing extends beyond immediate behavioural and perceptual benefits, and can also strongly impact future perception and behaviour. Some research investigated if benefits of multisensory memory extend to touch (Lehmann and Murray, 2005) and smell (Gottfried et al., 2004), but more systematic research is required to draw strong conclusions about the generalisability of the benefits of

multisensory memories across the different sensory systems. Likewise, it will be essential to equate baseline performance on tasks in all sensory modalities before more fully invoking inverse effectiveness as an explanation for larger memory benefits on the auditory than visual task in our studies.

### 3. Brain correlates of implicit multisensory benefits in memory

The majority of our brain mapping studies has focused on the networks involved in visual memory but all our studies employed the continuous old/new recognition paradigm described above (see Fig. 2b). In this section, we focus exclusively on brain responses elicited by repeated object presentations. Across both ERP and fMRI methods, portions of the lateral occipital cortex (LOC) were found to respond more strongly to naturalistic visual objects that had been initially accompanied by *semantically* congruent sounds (when compared to repeated images that were never presented together with sounds). More specifically, in the ERP study (Murray et al., 2004) we established that distinguishable brain networks (viz. ERP topographic differences) become active already within the first 60–135 ms post-stimulus (subsequent effects were also observed at ~210–260 ms and 318–390 ms). Source estimations identified the LOC as responding significantly more strongly towards visual objects previously presented in a semantically congruent multisensory context (see Fig. 3).

Subsequently, we have compared memory for objects presented only visually with that for objects presented initially in a *meaningless* multisensory context. That is, visual objects were all paired with the same, single tone (Lehmann and Murray, 2005) or each visual object was paired with a distinct tone (with tones modulated in their spectral composition, amplitude envelope and waveform type; Thelen et al., 2012). As highlighted above, such situations led to significant memory impairments when measured across the group of studied participants (Lehman and Murray, 2005; Thelen et al., 2012). In Thelen et al. (2012), ERP differences between objects initially presented exclusively visually and those presented in a multisensory manner (here with meaningless tones) began at ~100 ms post-stimulus. As in Murray et al. (2004), these effects were driven by changes in the ERP topography, rather than the strength of activation of the electric field at the scalp, suggesting that changes in the underlying sources were responsible for distinct memory performance. Source estimations of these differences were localised to a small cluster within the right LOC (as in Murray et al. (2004)) and a larger cluster in the right posterior superior temporal sulcus (pSTS). However, in Thelen et al. (2012) the LOC brain activity was weaker in the had-been multisensory when compared with the had-been unisensory condition, contrasting with the earlier study of Murray et al. (2004) where multisensory pairings had been semantically congruent. Responses in the pSTS were significantly stronger for the had-been multisensory than had-been unisensory condition. The brain distinguished between these two presentation types again at later stages (270–310 ms), with differences visible in the ERP topography and with source estimates in this time-period localised within the right middle temporal cortex. Notably, the strength of the activity within the right middle temporal cortex over the 270–310 ms period was directly related to the magnitude of performance impairment ( $r_{(10)} = 0.627$ ;  $p = 0.029$ ; cf. Fig. 3 in Thelen et al., 2012). Together, these findings suggest that the set of brain areas activated during the visual memory task is modulated by the effectiveness of a given multisensory encoding context on unisensory memory performance, rather than by the simple presence of the preceding multisensory context.

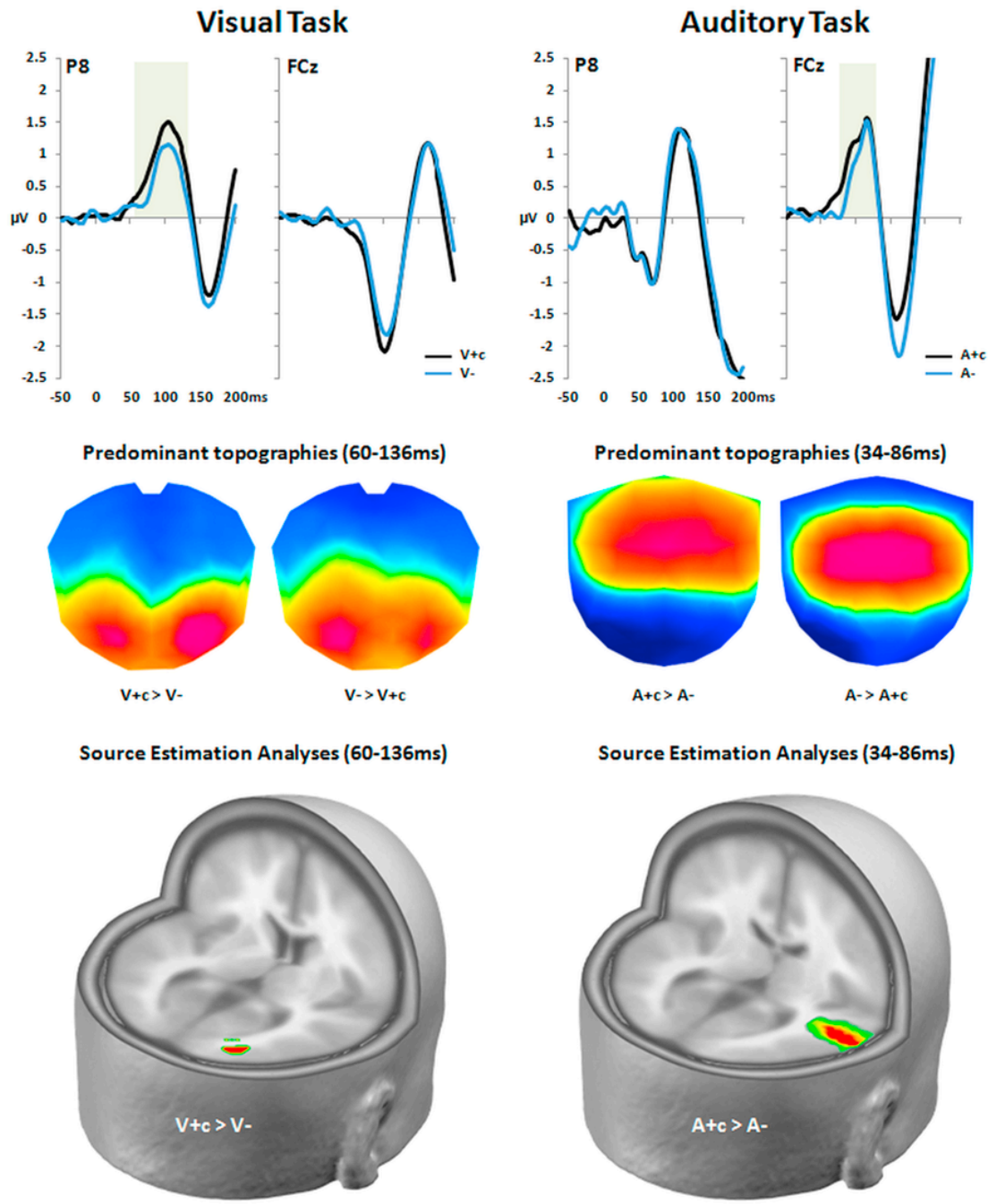
The benefits of multisensory processing for (episodic) memory for auditory objects seem to be supported by a relatively different set of brain areas and possibly altogether different brain mechanisms. In Matusz et al. (2015b), our participants discriminated naturalistic sounds that could be accompanied by semantically congruent images

or scrambled versions of these images and abstract figures. Notably, as in our purely behavioural study (Thelen et al., 2015), the multisensory benefits elicited in auditory memory by the initial presentation of objects in a semantically congruent multisensory manner were stronger than those for visual memory. We found that, yet again, the brain distinguished between objects based on their initial context at very early time points following stimulus presentation (i.e., 35–85 ms post-stimulus; see Fig. 3). However, the effect of initial context on auditory memory involved the right superior temporal cortex (rSTC), the right intraparietal cortex (rIPC), the right inferior occipital cortex, and left frontal cortex. Critically, the superior temporal and intraparietal cortices were those areas modulating in a manner that mapped on to changes in memory performance. Notably, the direction of these modulations was opposite to those found in the visual task in Murray et al. (2004). That is, the activity within these two brain areas was suppressed most strongly for multisensory congruent pairings compared to either auditory-only stimuli or multisensory meaningless pairings. This effect resembled the “response suppression” mechanism (Bergerbest et al., 2004; Murray et al., 2008) frequently reported as underlying short-lived learning effects within the auditory cortices. In Section 5, we discuss the implications of these findings for furthering our understanding of how objects are represented and accessed.

### 4. Individual differences in who benefits from multisensory contexts

Profound inter-individual differences were seen in our paradigm with healthy adults. When we analysed the results of our studies involving initial meaningless multisensory contexts in more detail (Thelen et al., 2014), a bimodal distribution of behavioural effects was observed. Specifically, a roughly equal proportion of participants improved as were impaired both when the task was visual and when it was auditory. Despite differences in timing, the same brain region (the intraparietal cortex; IPC) appeared to be a critical node in differentiating between individuals who were improved or impaired across both visual and auditory memory tasks. Importantly, there was no evidence for differences in how these groups of individuals processed unisensory, either visual or auditory, information. Current efforts by our lab are underway to better understand the nature of these differences. For example, do the groups differ because some individuals simply cannot help but integrate all multisensory events, while others are more capable of filtering out task-irrelevant information? It is important to point out that the groups did not differ in their overall performance (either accuracy or reaction time), which would run counter to an explanation based on differences in general distractibility. These points notwithstanding, the latency of the brain effects reported in Thelen et al. (2014) suggests that the underlying mechanism may be more reliant on how strongly multisensory simultaneity affected the selective attention of the observers towards the task-relevant unisensory stimuli. What requires further investigations is what specifically led IPC, a brain area well known to be involved in selective attention, as well as a critical hub for multisensory processing (Werner and Noppeney, 2010), to show stronger activity for those who benefited the most from the multisensory context of the initial stimulus encounter (and conversely showed weaker activity for those whose memory was impaired).

An important related domain of active inquiry by our group focuses on the hypothesis that an individual's capacity to integrate multisensory information, such as during a simple detection task, may directly scale to how an individual makes use of multisensory experiences to facilitate object recognition and memory. In other words, is one's ability to benefit from multisensory contexts during a memory task based on a more general capacity to integrate multisensory signals, such as simple beeps and flashes presented simultaneously at the same location? One limitation of our prior work is that all data came from different compo-



**Fig. 3.** Typical ERP findings showing differences between responses to unisensory stimuli (visual on the left side of the figure and auditory on the right side of the figure) according to whether they had been initially encountered in a semantically congruent multisensory context or unisensory context (V + c/A + c and V -/A -, respectively). The uppermost row shows ERPs from a right parieto-occipital electrode (P8) and fronto-central electrode (FCz). The shaded region shows periods of significant modulation. The middle row shows that these ERP modulations were due to topographic differences between conditions. Topographic maps are displayed on a flattened projection of the electrode montage, with nasion upward and the left hemisphere on the left. Red colours indicate positive potential, and blue colours indicate negative potential. The lowermost row shows loci of significant differences in distributed source estimations. For the visual task, stronger source activity was observed for V + c than V - within the right LOC. For the auditory task, stronger source activity was observed for A + c than A - within the right STC. Full details can be found in the original publications (Murray et al., 2004; Matusz et al., 2015b).

nents of the same paradigm (i.e. initial vs. repeated exposures during a continuous recognition task). Thus, it is unknown if links between multisensory processes persist when measured using two or more distinct tasks (each with their own stimulus set, goals, and attentional demands). Also, we do not know the extent to which any links in multisensory integrative capacity manifest specifically at behavioural and/or brain levels. Thus far, we have demonstrated a link between brain activity at one point in time and behaviour at a subsequent time point on the same task. Initial findings indeed point to links between behaviour

on a simple detection task and a standardized questionnaire indexing memory function (the mini mental state examination; Eardley et al., submitted).

**5. Cognitive mechanisms by which multisensory contexts improve memory**

Before we draw more general conclusions from our findings, we have to note that our paradigm investigates a very specific but etholog-

ically relevant situation. Namely, the task we utilise focuses on episodic memory (*Have you seen this object before in this experimental block?*), and the effectiveness of the processes underlying this memory system is investigated as a function of multisensory processes that are triggered likely outside of the observer's attentional focus (at least in many of our participants). Additionally, we have studied these processes predominantly with naturalistic objects (sounds and schematic drawings), thus likely triggering associations based on long-term, semantic associations between visual and auditory attributes of real-world objects, such as animals or tools (Beauchamp et al., 2004; Laurienti et al., 2004; Mahon and Caramazza, 2011; Werner and Noppeney, 2010). The nature of the processes engaged in our paradigm needs to be borne in mind when interpreting the observed brain and behavioural results and placing and interpreting them within the wider background of other studies on learning and memory in multisensory environments. This clarification helps to situate our findings within the broader literature of research on memory on the one hand and multisensory processing on the other.

One needs to distinguish our findings from those that have focused on the effortful encoding of components of new pairings (where the information presented in both sensory modalities is task-relevant; e.g. Nyberg et al., 2000; Wheeler et al., 2000; Von Kriegstein and Giraud, 2006). In these earlier works, there were discrepant findings regarding the benefit of semantically congruent multisensory contexts on later unisensory memory. By contrast, in our paradigm, the encoding process was done with the focus on a single task relevant sense. Thus the memory effects should have tapped, to some extent, bottom-up and stimulus-driven multisensory processes and as such should reveal the fuller impact of semantic congruence on memory. Here, presentations that engaged semantic memory improved episodic memory much more robustly than those that engaged processes triggered by audiovisual simultaneity detection alone. This pattern of results is in line with the benefit of activating other-modality representations of a given object within long-term semantic memory (Yuval-Greenberg and Deouell, 2007).

Despite these caveats, it is noteworthy that brain imaging and neurophysiology all demonstrate that responses to unisensory stimuli vary according to the context (i.e., unisensory vs. multisensory) in which they were either previously encountered or explicitly studied. What differs across these studies is whether the differential responses were indicative of a reactivation of a widespread network or were confined to modulations within a given brain region (Nyberg et al., 2000; Wheeler et al., 2000; Von Kriegstein and Giraud, 2006), albeit differing in its local pattern of activity (e.g., Gibson and Maunsell, 1997). Still other research would place a dominant role on medial temporal cortices, in particular perirhinal cortices, on the binding of semantic object features (Taylor et al., 2006; see also Murray and Bussey (1999)). Lesions to such cortices impaired performance on a delayed match-to-sample task, supporting a central role for this region in mediating encoding and retrieval processes that subservise (some forms of) multisensory memory. However, some of the discrepancies across studies may be explained by contributions of explicit attention to both sensory modalities as well as to the demands of the task (or not) for effortful encoding of the constituents of the multisensory pairings. In this regard, paradigms such as these may be building or accessing much richer representations than those at play in a unisensory continuous recognition paradigm. The latter instead promotes access to multisensory representations that is implicit, which is corroborated by the early and local nature of the modulations observed by us within task-relevant sensory cortices.

Specifically, our primary finding across both EEG and fMRI studies was that responses to repeated presentations of unisensory visual or auditory stimuli were affected implicitly and at early latencies by whether or not these stimuli had been previously presented in conjunc-

tion with a sound or image. This suggests that brain networks responsible for the processing of unisensory stimuli have access to multisensory memory representations early on in sensory-cognitive processing. Furthermore, our source estimations indicate that this access initially manifests within unisensory object recognition areas (as well as IPC in the case of the auditory memory task). We propose that this early modulation reflects the rapid reactivation of distinct multisensory and unisensory perceptual traces established during initial stimulus presentation. This notion is supported by findings from studies of repetition priming both in the visual (e.g., Doniger et al., 2001) as well as auditory modality (e.g. De Lucia et al., 2010; Murray et al., 2008), though we have no reason to suspect that repetition priming alone could account for our effects (cf. Murray et al., 2004 for a more extensive discussion). This proposal of distinct multisensory and unisensory perceptual traces is reinforced by two pieces of evidence: (1) that unisensory objection recognition areas demonstrate auditory-visual convergence, and (2) that multisensory memory representations are both localised and distinguishable from their unisensory counterparts. However, and despite evidence from single-unit intracranial recordings in non-human primates showing distinct representations for multisensory vs. unisensory paired associations (e.g. Gibson and Maunsell, 1997), we cannot fully discount the possibility (which may not be mutually exclusive with the above) that the initial multisensory experiences are instead impacting unisensory representations.

Nonetheless and first, it is now well established that visual areas such as the LOC and auditory areas such as the STC demonstrate multisensory convergence and integration (see, e.g., Matusz et al., 2016; Sarmiento et al., 2015; reviewed in Doehrmann and Naumer, (2008), Murray et al. (2016b), ten Oever et al. (2016)). Second, microelectrode recordings in monkey posterior infero-temporal (IT) cortex, for which the LOC is considered to be the human homologue, as well as visual area V4, demonstrate selective delay-period responses on a delayed match-to-sample task for specific multisensory and unisensory pairings (e.g., Colombo and Gross, 1994; Gibson and Maunsell, 1997; Haenny et al., 1988; Maunsell et al., 1991; see also Goulet and Murray (2001)). Neurons within these regions selectively distinguished unisensory stimuli according to their learned association with another stimulus of the same or different sensory modality. Crucially, the selective responses were specific to a given learned association; a neuron with multisensory selectivity did not also exhibit selectivity to other unisensory associations (Gibson and Maunsell, 1997). The implication is that there are distinct neural responses to and perhaps also distinct representations of unisensory and multisensory associations within patches of the IT cortex, which would satisfy the second prerequisite described above.

Our results extend this prior body of work by using task-irrelevant multisensory contexts. They demonstrate that the multisensory representations are 1) established within the cortices of the task-relevant sense and 2) are accessible subsequently. This combination in turn promotes stimulus discrimination during future unisensory stimulation. Thus, categorisation based on past experiences, at least in the early stages of brain processing, is supported by processes within the task-relevant cortices that themselves include multisensory representations. In the case of a visual discrimination task, for example, auditory cortices are not activated. These results contrast with some early hemodynamic results that demonstrated a close overlap of areas activated during memory encoding and retrieval (Nyberg et al., 2000; Wheeler et al., 2000). In this work, visual words learned as part of visual versus audio-visual pairs activated the same auditory areas as those involved when discriminated later on the basis of the sensory modality (or modalities). In the study of Nyberg et al. the absence of other activations may also be linked to the fact that their imaging results were masked by the results of the contrast between encoding visual and audio-visual conditions. Findings from these studies were taken as support for the "redintegration" hypothesis (Hamilton, 1859), according to



which the repeated presentation of stimuli in the task-relevant sense reactivates both the sensory-specific cortices (here, visual) as well as the secondary-modality cortices (here, auditory) because of a consolidated memory for the audio-visual association. However, it cannot be discounted that the additional activations in these studies were driven by participants using a form of auditory “imagery”, given they had to explicitly recall if the word had been originally learned with a sound. It should also be noted that performance was worse for visually-presented words previously learned with sounds.

The fact that our task does not require either effortful study of the multisensory pairs, or attention to the multisensory (or congruent) nature of the pairing, supports the interpretation of our results in terms of multisensory processes involuntarily creating (or accessing) distinct, durable representations for naturalistic images (as well as sounds) when accompanied by (in particular semantically congruent) stimuli in another sense. Direct comparison between our studies and this prior work is difficult based on the blocked nature of the study design and the poor temporal resolution of the prior work. Thus, reintegration processes may play a role in how memories are built and retrieved in multisensory settings. However, our results would necessarily constrain the nature of brain mechanisms governing these processes. First, reintegration may occur via multisensory processing, rather than via concomitant activity between the task-relevant sensory-specific auditory and visual cortices. Even then, our results constrain the possible time course of such effects. That is, if we look at the activations within the auditory cortex reported by Thelen et al. (2012), these were preceded by differential activity within generators localised within the LOC. Thus, any sensory-specific cortex activation would be expected to follow from reactivation of multisensory representations, rather than the initial auditory sensory experience. In other words, the repeated experience of the association may be a consequence, rather than a cause, of successful memory retrieval (see Rugg et al. (2008) for a similar critique of “reinstatement” accounts of memory retrieval brain mechanisms in visual memory research).

It is important to consider an alternative interpretation for the basis of these memory benefits, which is that multisensory exposures do not benefit memory processes per se, but rather only perception processes. Any situation, whether a multisensory exposure or some other emphasis of the task-relevant unisensory experience, would be sufficient to elicit the observed positive effects on later recognition memory, because the initial exposure is made more memorable in terms of its perception. While we cannot unequivocally exclude such an account at this stage, several aspects of our data would speak against a purely perceptual account.

First, on initial stimulus presentations there is no evidence for accuracy differences between unisensory and multisensory conditions (either semantically congruent, incongruent, or entailing a meaningless task-irrelevant stimulus). By contrast, reaction times for all multisensory stimuli were significantly slower than those to unisensory stimuli. This pattern was consistently observed both when the continuous recognition task was performed in the visual modality as well as when the task was performed in the auditory modality (cf. Fig. 2 in Thelen and Murray, 2013). If perceptual processes were mediating our effects, then a strong prediction would have been one of more accurate and faster behaviour to multisensory than unisensory stimuli during this initial encounter. Moreover, the similar pattern across all multisensory conditions on initial stimulus presentations does not explain the differential pattern on repeated presentations according to the semantic congruence of the initial multisensory exposure. Second, our ERP effects were consistent in terms of topographic modulations. This was the case both when the task was visual (Murray et al., 2004; Thelen et al., 2012) and when it was auditory (Matusz et al., 2015b). Because topographic modulations are forcibly the result of changes in the configuration of underlying brain sources, an explanation based on increased

salience or attention (i.e. typical perceptual processes) is unlikely as these have reliably led to changes in ERP amplitude or strength rather than topography. Finally, in addition to investigating ERP effects on stimulus repetitions, we have also reported how ERP responses during initial stimulus presentation are predictive of whether or not an individual's memory is enhanced upon repeated stimulus presentation (Thelen et al., 2014). Individuals who show memory enhancements also show stronger responses to initial multisensory stimulus presentations. This was not the case for unisensory stimulus presentations. An account where perceptual processes are the root of the enhanced memory would have predicted generally stronger responses (both multisensory and unisensory) in individuals exhibiting memory performance enhancement vs. those individual exhibiting memory performance decrements. This was not the case.

## 6. Broader implications

The demonstration of benefits from multisensory contexts on memory advances our understanding of both multisensory processes in general as well as of memory and the organisation of semantic knowledge.

First, in terms of implications for multisensory processing, the reviewed findings demonstrate that the products of multisensory processes persist over time, influencing subsequent unisensory object perception. Multisensory processes associated with the initial encounter of an object will influence the later retrieval of that object in an involuntary, incidental and general fashion. This underlines the importance of the individual's experience, both long-term as well as short-term (e.g., inter-trial effects), in influencing responses to both unisensory (visual, auditory, etc.) and multisensory objects. In agreement with these findings, recent theoretical frameworks have aimed to clarify the role of memory and its interplay with other top-down processes, such as selective attention, in controlling distinct multisensory processes (Murray et al., 2016; ten Oever et al., 2016; see also Matusz et al. (2015b) for evidence for impact of biases in selective attention towards threat-related visual stimuli on their subsequent memory). Understanding the dependencies of different multisensory processes on experience and attention has clear implications for supporting healthy learning as well as rehabilitation of sensory and learning disorders (Bach et al., 2015).

Second, these findings highlight the fact that the existing models of memory may not generalise to multisensory settings, particularly when notions such as conceptual novelty vs. physical familiarity come into play (e.g. Reggev et al., 2016). For one, the results reviewed here show that despite their task-irrelevance, semantically congruent contexts bring benefits over and above those predicted by some of the most fundamental hypotheses developed within visual memory research, such as the encoding-retrieval congruence hypothesis (Baddeley et al., 2009; Smith and Vela, 2001). As such, these results bridge traditionally separate lines of research by demonstrating that memory processes may be coupled with more general multisensory processes, such as those readily studied in simple, detection-like perceptual tasks. Further research will be required to ascertain the extent to which the capacity to benefit from multisensory contexts for memory functions is yoked upon a more general capability to benefit from multisensory information in the environment. Ongoing efforts from our group are investigating the extent to which both school-aged children as well as the elderly exhibit such links.

Furthermore, our findings provide direct evidence to the models of functional brain organisation that propose that object representations are quintessentially multisensory. In agreement, recent conceptual frameworks suggest that higher-order visual cortices may be better defined by the tasks they subserve than by the sensory inputs they are traditionally thought to receive (Murray et al., 2016; Reich et al., 2012; ten Oever et al., 2016). In other words, some regions may be special-

ized to perform a given object recognition function (e.g. face, letters, body parts, etc.), irrespective of whether object-relevant information is conveyed by images, sounds, or touch. This has been demonstrated both when environmental sounds of the objects were presented as well as when same objects were re-coded via sensory substitution devices. For example, Mahon et al. (2009) show that the topological pattern observed in the sighted with visual stimuli was the same as that in the congenitally blind when sounds were used. While memory functions were not explicitly taxed in the majority of these studies, they undoubtedly contribute to the establishment and maintenance of these representations (e.g., Amedi et al., 2003). In this perspective, our results enrich our understanding of how these multisensory representations are accessed even in unisensory contexts. We would underscore that sensory modality does indeed play an important role in memory. It does so to the extent that it determines the cortices (and likely also the mechanisms) that will support the retrieval of multisensory memories. Likewise, these multisensory contexts provide an implicit “tag” to experienced events that can facilitate subsequent recognition. We are currently investigating what limitations apply to this tagging and by extension their behavioural consequences. Finally, the evidence thus far would indicate that semantic, memory-based processes seem to be most effective in conferring benefits upon memory, while processes based on low-level factors such as stimulus timing appear to exhibit higher degrees of inter-individual variability. Therefore, it will be crucial to determine how to optimize these functions at the service of learning and memory.

#### Uncited references

Bein et al. (2015), McAndrews et al. (2016), Murray et al. (2015).

#### Acknowledgements

We thank Antonia Thelen for comments on the initial drafts of the manuscript. PJM receives support from the Pierre Mercier Foundation. MTW receives support from the National Institutes of Health (grants CA183492, DC014114, HD083211, MH109225), from the Simons Foundation Autism Research Initiative and from the Wallace Foundation. MMM receives support from the Swiss National Science Foundation (grants 320030-149982 and 320030-169206 as well as National Centre of Competence in Research project “SYNAPSY, The Synaptic Bases of Mental Disease” [project 51AU40-125759]) and Carigest.

#### References

- Amedi, A., Raz, N., Pianka, P., Malach, R., Zohary, E., 2003. Early “visual” cortex activation correlates with superior verbal memory performance in the blind. *Nat. Neurosci.* 6 (7), 758–766. <http://dx.doi.org/10.1038/nn1072>.
- Bach, S., Richardson, U., Brandeis, D., Martin, E., Brem, S., 2013. Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *Neuroimage* 82, 605–615.
- Baddeley, A., Eysenck, M.W., Anderson, M.C., 2009. *Memory*. Psychology Press.
- Baum, S.H., Stevenson, R.A., Wallace, M.T., 2015. Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Prog. Neurobiol.* 134, 140–160.
- Beauchamp, M.S., Argall, B.D., Bodurka, J., Duyn, J.H., Martin, A., 2004. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat. Neurosci.* 7 (11), 1190–1192. <http://dx.doi.org/10.1038/nn1333>.
- Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., Maril, A., 2015. Delineating the effect of semantic congruency on episodic memory: the role of integration and relatedness. *PLoS One* 10 (2), 1–24. <http://dx.doi.org/10.1371/journal.pone.0115624>.
- Bergerbest, D., Ghahremani, D.G., Gabrieli, J.D.E., 2004. Neural correlates of auditory repetition priming: reduced fMRI activation in the auditory cortex. *J. Cogn. Neurosci.* 16 (6), 966–977.
- Cohen, M.A., Horowitz, T.S., Wolfe, J.M., 2009. Auditory recognition memory is inferior to visual recognition memory. *Proc. Natl. Acad. Sci. USA* 106 (14), 6008–6010. <http://dx.doi.org/10.1073/pnas.0811884106>.

- Colombo, M., Gross, C.G., 1994. Responses of inferior temporal cortex and hippocampal neurons during delayed matching-to-sample in monkeys (*Macaca fascicularis*). *Behav. Neurosci.* 108, 443–455.
- Constantinescu, A.O., O’Reilly, J.X., Behrens, T.E.J., 2016. Organizing conceptual knowledge in humans with a gridlike code. *Science* 352 (6292), 1464–1468.
- Crosse, M.J., Di Liberto, G.M., Lalor, E.C., 2016. Eye can hear clearly now: inverse effectiveness in natural audiovisual speech processing relies on long-term crossmodal temporal integration. *J. Neurosci.* 36 (38), 9888–9895.
- De Lucia, M., Cocchi, L., Martuzzi, R., Meuli, R.A., Clarke, S., Murray, M.M., 2010. Perceptual and semantic contributions to repetition priming of environmental sounds. *Cereb. Cortex* 20 (7), 1676–1684. <http://dx.doi.org/10.1093/cercor/bhp230>.
- De Meo, R., Murray, M.M., Clarke, S., Matusz, P.J., 2015. Top-down control and early multisensory processes: chicken vs. egg. *Front. Integr. Neurosci.* 9, 1–6. <http://dx.doi.org/10.3389/fnint.2015.00017>.
- Doehrmann, O., Naumer, M.J., 2008. Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain Res.* 1242, 136–150. <http://dx.doi.org/10.1016/j.brainres.2008.03.071>.
- Doniger, G.M., Foxe, J.J., Schroeder, C.E., Murray, M.M., Higgins, B.A., Javitt, D.C., 2001. Visual perceptual learning in human object recognition areas: a repetition priming study using high-density electrical mapping. *NeuroImage* 13 (2), 305–313. <http://dx.doi.org/10.1006/nimg.2000.0684>.
- Eardley, A.F., Edginton, T., Oyekan, R., Smyth, E., Murray, M.M., Matusz, P.J., 2017. Linking multisensory and memory functions in ageing. *Neurobiol. Aging* (in preparation).
- Gazzaley, A., Nobre, A.C., 2012. Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.* <http://dx.doi.org/10.1016/j.tics.2011.11.014>.
- Gibson, J.R., Maunsell, J.H.R., 1997. Sensory modality specificity of neural activity related to memory in visual cortex. *J. Neurophysiol.* 78, 1263–1275.
- Gingras, G., Rowland, B.A., Stien, B.E., 2009. The differing impact of multisensory and unisensory integration on behavior. *J. Neurosci.* 29, 4897–4902.
- Gottfried, J.A., Smith, A.P.R., Rugg, M.D., Dolan, R.J., 2004. Remembrance of odors past: human olfactory cortex in cross-modal recognition memory. *Neuron* 42 (4), 687–695. [http://dx.doi.org/10.1016/S0896-6273\(04\)00270-3](http://dx.doi.org/10.1016/S0896-6273(04)00270-3).
- Goulet, S., Murray, E.A., 2001. Neural substrates of crossmodal association memory in monkeys: the amygdala versus the anterior rhinal cortex. *Behav. Neurosci.* 115 (2), 271.
- Haenny, P.E., Maunsell, J.H.R., Schiller, P.H., 1988. State dependent activity in monkey visual cortex II. Retinal and extraretinal factors in V4. *Exp. Brain Res.* 69, 245–259.
- Hamilton, W., 1859. *Lectures on Metaphysics and Logic*, 1. Gould and Lincoln, Boston, MA (USA).
- Heikkilä, J., Alho, K., Hyvönen, H., Tiippana, K., 2015. Audiovisual semantic congruency during encoding enhances memory performance. *Exp. Psychol.* 62 (2), 123–130. <http://dx.doi.org/10.1027/1618-3169/a000279>.
- Heikkilä, J., Tiippana, K., 2016. School-aged children can benefit from audiovisual semantic congruency during memory encoding. *Exp. Brain Res.* 234 (5), 1199–1207. <http://dx.doi.org/10.1007/s00221-015-4341-6>.
- Jiang, W., Wallace, M.T., Jiang, H., Vaughan, J.W., Stein, B.E., 2001. Two cortical areas mediate multisensory integration in superior colliculus neurons. *J. Neurophysiol.* 85, 506–522.
- Johansson, B.B., 2012. Multisensory stimulation in stroke rehabilitation. *Front. Hum. Neurosci.* 6, 60.
- Laurienti, P.J., Kraft, R.A., Maldjian, J.A., Burdette, J.H., Wallace, M.T., 2004. Semantic congruence is a critical factor in multisensory behavioral performance. *Exp. Brain Res.* 158 (4), 405–414. <http://dx.doi.org/10.1007/s00221-004-1913-2>.
- Lehmann, S., Murray, M.M., 2005. The role of multisensory memories in unisensory object discrimination. *Cogn. Brain Res.* 24 (2), 326–334. <http://dx.doi.org/10.1016/j.cogbrainres.2005.02.005>.
- Lewis, J.W., 2010. Audio-visual perception of everyday natural object - hemodynamic studies in humans. In: Naumer, Kaiser (Eds.), *Multisensory Object Perception in the Primate Brain*. Springer, New York (USA).
- Macmillan, N.A., Creelman, C.D., 2004. *Detection Theory: A User’s Guide*. Psychology press.
- Mahon, B.Z., Caramazza, A., 2011. What drives the organization of object knowledge in the brain? The distributed domain-specific hypothesis. *Trends Cogn. Sci.* 15, 97–103. <http://dx.doi.org/10.1016/j.tics.2011.01.004>.
- Matusz, P.J., Thelen, A., Amrein, S., Geiser, E., Anken, J., Murray, M.M., 2015. The role of auditory cortices in the retrieval of single-trial auditory-visual object memories. *Eur. J. Neurosci.* <http://dx.doi.org/10.1111/ejn.12804>.
- Matusz, P.J., Retsa, C., Murray, M.M., 2016. The context-contingent nature of cross-modal activations of the visual cortex. *NeuroImage* 125, 996–1004.
- Matusz, P.J., Traczyk, J., Sobkow, A., Strelau, J., 2015. Individual differences in emotional reactivity moderate the strength of the relationship between attentional and implicit-memory biases towards threat-related stimuli. *J. Cogn. Psychol.* <http://dx.doi.org/10.1080/20445911.2015.1027210>.
- Maunsell, J.H.R., Sclar, G., Nealey, T.A., DePriest, D.D., 1991. Extraretinal representations in area V4 in the macaque monkey. *Vis. Neurosci.* 7 (6), 561–573. <http://dx.doi.org/10.1017/S095252380001035X>.
- McAndrews, M.P., Girard, T.A., Wilkins, L.K., McCormick, C., 2016. Semantic congruence affects hippocampal response to repetition of visual associations. *Neuropsychologia* 90, 235–242. <http://dx.doi.org/10.1016/j.neuropsychologia.2016.07.026>.
- Moran, Z.D., Bachman, P., Pham, P., Hah Cho, S., Cannon, T.D., Shams, L., 2013. Multisensory encoding improves auditory recognition. *Multisens. Res.* 26, 581–592. <http://dx.doi.org/10.1163/22134808-00002436>.

- Murray, E.A., Bussey, T.J., 1999. Perceptual-mnemonic functions of the perirhinal cortex. *Trends Cogn. Sci.* 3 (4), 142–151. [http://dx.doi.org/10.1016/S1364-6613\(99\)01303-0](http://dx.doi.org/10.1016/S1364-6613(99)01303-0).
- Murray, M.M., Michel, C.M., De Peralta, R.G., Ortigue, S., Brunet, D., Andino, S.G., Schneider, A., 2004. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage* 21 (1), 125–135. <http://dx.doi.org/10.1016/j.neuroimage.2003.09.035>.
- 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. <http://dx.doi.org/10.1016/j.neuroimage.2005.04.016>.
- Murray, M.M., Camen, C., Spierer, L., Clarke, S., 2008. Plasticity in representations of environmental sounds revealed by electrical neuroimaging. *NeuroImage* 39 (2), 847–856.
- Murray, M.M., Wallace, M.T., 2012. *The Neural Bases of Multisensory Processes*. CRC Press, Boca Raton, FL (USA).
- Murray, M.M., Matusz, P.J., Amedi, A., 2015. Neuroplasticity: unexpected consequences of early blindness. *Curr. Biol.* 25 (20), R998–R1001. <http://dx.doi.org/10.1016/j.cub.2015.08.054>.
- Murray, M.M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., Matusz, P.J., 2016. The multisensory function of primary visual cortex in humans. *Neuropsychologia* (Retrieved from) (<http://www.sciencedirect.com/science/article/pii/S0028393215301275>).
- Murray, M.M., Lewkowicz, D.J., Amedi, A., Wallace, M.T., 2016. Multisensory processes: a balancing act across the lifespan. *Trends Neurosci.* 39 (8), 567–579. <http://dx.doi.org/10.1016/j.tins.2016.05.003>.
- 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. <http://dx.doi.org/10.1016/j.neulet.2011.10.022>.
- Nyberg, L., Habib, R., McIntosh, A.R., Tulving, E., 2000. Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl. Acad. Sci. USA* 97, 11120–11124. <http://dx.doi.org/10.1073/pnas.97.20.11120>.
- Reggev, N., Bein, O., Maril, A., 2016. Distinct neural suppression and encoding effects for conceptual novelty and familiarity. *J. Cogn. Neurosci.* Cogn. Neurosci. 28, 1455–1470. <http://dx.doi.org/10.1162/jocn.2016.03.001>.
- Reich, L., Maidenbaum, S., Amedi, A., 2012. The brain as a flexible task machine: implications for visual rehabilitation using noninvasive vs. invasive approaches. *Curr. Opin. Neurol.* 25 (1), 86–95. <http://dx.doi.org/10.1097/WCO.0b013e32834ed723>.
- Romei, V., Murray, M.M., Cappe, C., Thut, G., 2009. Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Curr. Biol.* 19 (21), 1799–1805. <http://dx.doi.org/10.1016/j.cub.2009.09.027>.
- Romei, V., Murray, M.M., Cappe, C., Thut, G., 2013. The contributions of sensory dominance and attentional bias to cross-modal enhancement of visual cortex excitability. *J. Cogn. Neurosci.* 25, 1122–1135. <http://dx.doi.org/10.1162/jocn.2013.01.001>.
- 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. [http://dx.doi.org/10.1016/S0079-6123\(07\)00021-0](http://dx.doi.org/10.1016/S0079-6123(07)00021-0).
- Sarmiento, B.R., Matusz, P.J., Sanabria, D., Murray, M.M., 2015. Contextual factors multiplex to control multisensory processes. *Hum. Brain Mapp.* 37, 273–288. <http://dx.doi.org/10.1002/hbm.23030>.
- Schmid, C., Büchel, C., Rose, M., 2011. The neural basis of visual dominance in the context of audio-visual object processing. *NeuroImage* 55 (1), 304–311. <http://dx.doi.org/10.1016/j.neuroimage.2010.11.051>.
- 2001 S.M. Smith, E. Vela, Environmental context-dependent memory, *Psychon. Bull. Rev.* 8 (2001) 203–220. <http://dx.doi.org/10.3758/BF03196157>.
- Stevenson, R.A., Ghose, D., Fister, J.K., Sarko, D.K., Altieri, N.A., Nidiffer, A.R., Wallace, M.T., 2014. Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr.* 27 (6), 707–730. <http://dx.doi.org/10.1007/s10548-014-0365-7>.
- Stevenson, R.A., James, T.W., 2009. Audiovisual integration in human superior temporal sulcus: inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage* 44 (3), 1210–1223. <http://dx.doi.org/10.1016/j.neuroimage.2008.09.034>.
- Summerfield, C., de Lange, F.P., 2014. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 15, 745–756. <http://dx.doi.org/10.1038/nrn3838>.
- Taylor, K.I., Moss, H.E., Stamatakis, E.A., Tyler, L.K., 2006. Binding crossmodal object features in perirhinal cortex. *Proc. Natl. Acad. Sci. USA* 103 (21), 8239–8244. <http://dx.doi.org/10.1073/pnas.0509704103>.
- ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M.M., Matusz, P.J., 2016. The COGs (context, object, goals) in multisensory processing. *Exp. Brain Res.* 234, 1307–1323. <http://dx.doi.org/10.1007/s00221-016-4590-z>.
- Thelen, A., Cappe, C., Murray, M.M., 2012. Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *NeuroImage* 62, 1478–1488. <http://dx.doi.org/10.1016/j.neuroimage.2012.05.027>.
- Thelen, A., Matusz, P.J., Murray, M.M., 2014. Multisensory context portends object memory. *Curr. Biol.* 24, R434–5. (<http://www.sciencedirect.com/science/article/pii/S0960982214007568>).
- Thelen, A., Murray, M.M., 2013. The efficacy of single-trial multisensory memories. *Multisens. Res.* 26 (5), 483–502. <http://dx.doi.org/10.1163/22134808-00002426>.
- Thelen, A., Talsma, D., Murray, M.M., 2015. Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition* 138, 148–160. <http://dx.doi.org/10.1016/j.cognition.2015.02.003>.
- Ueno, D., Masumoto, K., Sutani, K., Iwaki, S., 2015. Latency of modality-specific reactivation of auditory and visual information during episodic memory retrieval. *Neuroreport* 26 (6), 303–308. <http://dx.doi.org/10.1097/WNR.0000000000000325>.
- Von Kriegstein, K., Giraud, A.L., 2006. Implicit multisensory associations influence voice recognition. *PLoS Biol.* 4 (10), 1809–1820. <http://dx.doi.org/10.1371/journal.pbio.0040326>.
- Wallace, M.T., 2004. The development of multisensory processes. *Cogn. Process.* 5 (2), 69–83. <http://dx.doi.org/10.1007/s10339-004-0017-z>.
- Wallace, M.T., Carriere, B.N., Perrault, T.J., Vaughan, J.W., Stein, B.E., 2006. The development of cortical multisensory integration. *J. Neurosci.* 26 (46), 11844–11849. <http://dx.doi.org/10.1523/JNEUROSCI.3295-06.2006>.
- Wallace, M.T., Meredith, M.A., Stein, B.E., 1998. Multisensory integration in the superior colliculus of the alert cat. *J. Neurophysiol.* 80, 1006–1010.
- Wallace, M.T., Perrault, T.J., Hairston, W.D., Stein, B.E., 2004. Visual experience is necessary for the development of multisensory integration. *J. Neurosci.* 24 (43), 9580–9584. <http://dx.doi.org/10.1523/JNEUROSCI.2535-04.2004>.
- 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. <http://dx.doi.org/10.1523/JNEUROSCI.5091-09.2010>.
- Wheeler, M.E., Petersen, S.E., Buckner, R.L., 2000. Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc. Natl. Acad. Sci. USA* 97, 11125–11129.
- Yuval-Greenberg, S., Deouell, L.Y., 2007. What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. *J. Neurosci.* 27 (5), 1090–1096. <http://dx.doi.org/10.1523/jneurosci.4828-06.2007>.
- Yuval-Greenberg, S., Deouell, L.Y., 2009. The dog's meow asymmetrical interaction in cross-modal object recognition. *Exp. Brain Res.* 193 (4), 603–614.