

Serveur Académique Lausannois SERVAL serval.unil.ch

Author Manuscript

Faculty of Biology and Medicine Publication

This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Published in final edited form as:

Title: A multisensory perspective on object memory.

Authors: Matusz PJ, Wallace MT, Murray MM

Journal: Neuropsychologia

Year: 2017 Apr 8

DOI: [10.1016/j.neuropsychologia.2017.04.008](https://doi.org/10.1016/j.neuropsychologia.2017.04.008)

In the absence of a copyright statement, users should assume that standard copyright protection applies, unless the article contains an explicit statement to the contrary. In case of doubt, contact the journal publisher to verify the copyright status of an article.

A multisensory perspective on object memory

Pawel J. Matusz¹, Mark T. Wallace²⁻⁶, Micah M. Murray^{1,2,7,8,*}

¹The Laboratory for Investigative Neurophysiology (The LINE), Neuropsychology & Neurorehabilitation Service & Department of Radiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland

²Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

³Department of Psychology, Vanderbilt University, Nashville, TN, USA

⁴Kennedy Center for Research on Human Development, Vanderbilt University, Nashville, TN, USA

⁵Department of Psychiatry, Vanderbilt University, Nashville, TN, USA

⁶Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA

⁷EEG Brain Mapping Core, Center for Biomedical Imaging (CIBM) of Lausanne and Geneva, Lausanne, Switzerland

⁸Department of Ophthalmology, University of Lausanne, Jules-Gonin Eye Hospital, Lausanne, Switzerland

*Corresponding Author

Micah.murray@chuv.ch

Abstract

1
2
3
4 Traditional studies of memory and object recognition involved objects presented within a single
5
6 sensory modality (i.e., purely visual or purely auditory objects). However, in naturalistic settings,
7
8 objects are often evaluated and processed in a multisensory manner. This begets the question of
9
10 how object representations that combine information from the different senses are created and
11
12 utilised by memory functions. Here we review research that has demonstrated that a single
13
14 multisensory exposure can influence memory for both visual and auditory objects. In an old/new
15
16 object discrimination task, objects that were presented initially with a task-irrelevant stimulus in
17
18 another sense were better remembered compared to stimuli presented alone, most notably when
19
20 the two stimuli were semantically congruent. The brain discriminates between these two types of
21
22 object representations within the first 100ms post-stimulus onset, indicating early “tagging” of
23
24 objects/events by the brain based on the nature of their initial presentation context. Interestingly,
25
26 the specific brain networks supporting the improved object recognition vary based on a variety of
27
28 factors, including the effectiveness of the initial multisensory presentation and the sense that is task-
29
30 relevant. We specify the requisite conditions for multisensory contexts to improve object
31
32 discrimination following single exposures, and the individual differences that exist with respect to
33
34 these improvements. Our results shed light onto how memory operates on the multisensory nature
35
36 of object representations as well as how the brain stores and retrieves memories of objects.
37
38
39
40
41
42
43
44
45
46
47
48

Key words:

49
50 Multisensory, cross-modal, auditory, visual, object, memory, learning
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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 (Figure 1)¹. 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, O'Reilly, & Behrens, 2016; Gazzaley & Nobre, 2012; Mahon & Caramazza, 2011; Summerfield & 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 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 recognize Sarah better than her friends on your second meeting in the example provided above. In the current

¹Please 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.

1 review, we discuss the evidence that multisensory context can improve unisensory object
2 discrimination even after a single exposure. We then specify the requisite conditions for such
3 improvements as well as the individual differences therein. Lastly, we place the reviewed
4 behavioural and brain imaging findings within the broader literature on multisensory learning and
5 discuss the importance of considering multisensory contributions when creating accurate models of
6 object perception and memory.
7
8
9
10
11
12

13 While the experimental paradigm that we have employed has been described in detail
14 previously (Thelen & Murray, 2013; for a summary, see [Figure 2a](#)), we summarise it here briefly. We
15 employed a continuous recognition task, in which on each trial participants have to indicate as
16 quickly and accurately as possible whether they saw a given object for the first (“new”) or second
17 (“old”) time. Across different variations of this paradigm utilised in a number of studies over the
18 years, stimuli within one sense (e.g. vision) would always be task-relevant, while stimuli in another
19 sense (e.g. audition) would always be task-irrelevant. The initial and repeated trials were always
20 equally probable, and across all trial types the number of unisensory and multisensory trials were
21 also equally probably distributed. While it was the case that some of our early work involved a
22 paradigm where multisensory information was only presented on initial trials, subsequent work has
23 replicated effects even when rendering the multisensory content uninformative about the task-
24 relevant dimension (i.e. whether an object was presented for the initial or repeated time). In this line
25 of research, the effectiveness of three distinct multisensory contexts in improving memory has been
26 assessed: 1) a semantically congruent context – where the task-relevant and task-irrelevant stimuli
27 represent the same object (e.g., a drawing of a cow combined with a sound “moo”), 2) a
28 meaningless-association context – where the task-relevant stimuli are paired with tones or noises,
29 and 3) a semantically incongruent context – where the task-relevant and task-irrelevant stimuli
30 represent different objects.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

56 By manipulating the number (and type) of senses actively engaged, the nature of the
57 relationship between the stimuli across the two senses, as well as their task-relevance, our paradigm
58
59
60
61
62
63
64
65

1 sought to more closely emulate information processing in naturalistic environments. This evidence
2 (and more recently that from other independent laboratories) has provided novel insights into the
3 behavioural and brain mechanisms guiding memory and information processing in everyday
4 situations. The overall message from these studies is that memory for objects is generally improved
5 when the information is first encountered in a multisensory manner.
6
7
8
9

10 11 12 13 14 **1. Which multisensory contexts improve memory?** 15 16 17

18
19 In our paradigm, the benefits on object memory of having information presented in a
20 multisensory manner are generally observed as improved discrimination accuracy. Reaction times
21 (RTs) showed no similar benefits (e.g., Lehmann & Murray, 2005). When the initial multisensory
22 presentation (and encoding) involved semantically congruent pairings, robust memory
23 improvements were observed on subsequent retrieval. These improvements were observed across
24 studies employing different stimulus and paradigm parameters that balanced the occurrence of
25 multisensory information on initial and repeated presentations, and distinct brain mapping methods.
26 For example, these improvements in discrimination were seen in Murray et al.'s (2005) fMRI study,
27 despite the presence of scanner noise that arguably could have interfered with the ability to
28 perceive/encode the task-irrelevant sounds and also extended the usual item repetition lag of 5
29 seconds (used in the psychophysical and EEG studies) up to 50 seconds to accommodate inter-trial
30 intervals necessary due to the constraints of fMRI data acquisition (see Table 1). Across studies, the
31 observed benefits for semantically congruent multisensory pairings on memory performance ranged
32 from a gain of 2.5% to 9% over performance on unisensory visual or auditory trials (also [Figure 2b](#)).
33 Where the study design enabled the calculation of a more rigorous measure of sensory processing
34 (the perceptual sensitivity parameter, d' , Macmillan & Creelman, 2004), these multisensory benefits
35 were found to be even larger (i.e., 12% performance memory improvement; Matusz et al., 2015).
36 Overall, these improvements have been seen across 6 studies involving more than 100 participants
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 and exhibiting effect sizes ranging from small to large ($\eta^2_p = 0.14-0.63$; see Table 1 of Thelen and
2 Murray, 2013 for details; see also Moran et al., 2013; for similar size of effects in studies involving
3
4 setups with separate exposure and recall, see Heikkilä, Alho, Hyvönen, & Tiippana, 2015; Heikkilä &
5
6 Tiippana, 2016; Naghavi, Eriksson, Larsson, & Nyberg, 2011; Ueno, Masumoto, Sutani, & Iwaki,
7
8 2015).

9
10
11 In contrast to when multisensory stimuli were semantically congruent, if the initial pairing is
12
13 semantically incongruent, the typical result is memory *impairments* relative to when stimuli are
14
15 initially presented in a unisensory manner, with the impairments ranging from between a 4% and
16
17 16.5% decrease in discrimination accuracy (Figure 2b). Similarly, if the initial presentations involved
18
19 pairings with meaningless task-irrelevant information, performance decrements of 3-4% were
20
21 typically seen (Figure 2b). Intriguingly, for these meaningless pairings, performance was highly
22
23 variable across individuals. Thus, approximately half of the tested participants demonstrated
24
25 memory improvements following these meaningless initial contexts, and these improvements were
26
27 seen for both visual and auditory memory². For visual memory, the improvements ranged between
28
29 0.5% and 7% (Thelen, Matusz, & Murray, 2014), while for auditory memory these gains were
30
31 between 2.5% and 10.8% (Thelen et al. 2014, Supplemental Information). These findings suggest
32
33 that multisensory pairings involving merely the simultaneity of stimulus onsets across the senses
34
35 may only be weakly effective in supporting object memory (cf., De Meo, Murray, Clarke, & Matusz,
36
37 2015; Murray et al., 2016a,b; ten Oever et al., 2016 for reviews on the role of audiovisual
38
39 simultaneity detection in modulating instantaneous perception and selective attention). We discuss
40
41 these results in more detail below in the section on individual differences (Section 3).
42
43
44
45
46
47
48

49 To summarise, in a continuous discrimination paradigm, the initial presentation of stimuli in a
50
51 multisensory context (as opposed to unisensory), whether congruent or incongruent, has significant
52
53 influence on memory performance. Across studies, we have identified several requisite conditions
54
55 under which initial multisensory presentation improve memory (Table 1). These conditions, and the
56
57

58
59 ² We would note that a similar degree of inter-individual variability is observed with semantically incongruent
60 pairings, though to date this has not been specifically investigated with brain mapping/imaging methods.
61

1 effects that we observed as they were manipulated, challenge some of the basic tenets of traditional
2 models of memory. First, one of the most established findings in this domain is that memory
3 performance is best in situations where the encoding and retrieval contexts are identical (Baddeley,
4 Eysenck, & Anderson, 2009; Smith & Vela, 2001). Our results strongly suggest that those findings
5 hold true mainly, if not exclusively, in unisensory settings. Naturalistic environments, whether it be a
6 classroom or a cocktail party, are typically multisensory in nature. According to our results, in such
7 naturalistic environments, any memory benefits can be further enhanced in cases where the
8 encoded object stimulated multiple senses. If the signal in the other sense semantically matches the
9 identity of the object presented in the task-relevant sense and, thus, activate presumably long-term
10 memory associations, these benefits will likely be visible in all individuals. We would emphasize that
11 attention was always focused on one sense exclusively. Yet, the benefits of multisensory
12 presentations were nonetheless observed and thus can be considered implicit. This highlights the
13 efficacy of multisensory processes in influencing object memory even in situations where they occur
14 outside of the focus of selective attention.

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33 An important issue is to what extent these memory benefits are specific to multisensory
34 presentations. To our knowledge, this has not yet been specifically investigated with a continuous
35 recognition paradigm similar to that which we have used. That said, there are behavioural data in
36 cats showing that performance enhancements are greater for multisensory than for unisensory
37 redundancy (Gingras, Rowland, & Stein, 2009). Likewise, multisensory benefits were found to be
38 greater than unisensory benefits during a masked letter identification task in humans (Chen &
39 Spence, 2011). Finally and most germane are data from two studies. One study had non-human
40 primates performing a delayed match-to-sample task (Gibson & Maunsell, 1997). They showed that
41 the propensity of selective delay period activity at IT neurons was significantly greater for
42 multisensory than for unisensory learned associations. The other study had human participants
43 perform a recognition memory task with separate learning and memory phases (Heikkila et al.,
44 2015). Semantic congruence at encoding was beneficial for later recognition memory when the

1 materials were multisensory, but not when they were unisensory visual. Collectively these results,
2 alongside our own data, would support an account based on multisensory processing.
3

4 Multisensory interactions based on temporal relations (i.e., onset simultaneity) or on well-
5 learned associations can influence memory despite the multisensory aspect being task-irrelevant
6 and thus outside the goals of the observer. The presence of these multisensory-based memory
7 benefits has important clinical implications for treatment and rehabilitation of memory and sensory
8 disorders (e.g. Johansson, 2012; Baum, Stevenson & Wallace 2015), and, thus, an important question
9 is how generalisable these results are. Our ongoing work, which demonstrated a link between the
10 strength of multisensory benefits on a simple reaction time task and a degree of preservation of
11 higher-level functioning assessed with a standardised questionnaire in individuals with a mild
12 cognitive impairment, is focusing specifically on this question (Eardley et al., in review). While we
13 discuss the prerequisite conditions for these memory modulations to occur in Section 3, we will now
14 discuss how the observed memory effects change as a function of the task-relevant sense.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

30 The last several years have seen a growing interest in the role of the task-relevant sense as a
31 bottom-up factor modulating multisensory processing (Romei et al., 2009, 2013; Schmid et al.,
32 2011). Auditory object memory is generally reported to be weaker than visual memory (Cohen et al.,
33 2009; Yuval-Greenberg & Deouell, 2007, 2009). This poses an important question of whether
34 benefits from encoding stimuli in a multisensory context would be larger for auditory than visual
35 memory. Such a finding would be similar to observations that stronger benefits of multisensory
36 processing are frequently observed in situations in which the inputs are weakly effective (“inverse
37 effectiveness principle”, Crosse, Di Liberto, & Lalor, 2016; Jiang, Wallace, Jiang, Vaughan, & Stein,
38 2001; Stevenson et al., 2014; Stevenson & James, 2009; Wallace, 2004; Wallace, Carriere, Perrault,
39 Vaughan, & Stein, 2006; Wallace, Meredith, & Stein, 1998; Wallace, Perrault, Hairston, & Stein,
40 2004). Evidence in support of this notion was found in one of our studies (Thelen et al., 2015), where
41 the same group of participants performed the old/new task first with visual and then with auditory
42 objects as task-relevant objects, or vice versa. As in other studies (Cohen et al., 2009; Yuval-
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 Greenberg & Deouell, 2007, 2009), auditory memory was generally much weaker than visual
2 memory (67% vs. 92% accuracy, respectively). Similarly to our previous studies, we found memory
3 benefits that were exclusive for semantically congruent pairings, and that were seen for both vision
4 and hearing. Notably, the auditory memory benefits were approximately four times larger than the
5 visual benefits (i.e., 8.8% vs. 2.2% improvement in object discrimination). Because a very similar
6 setup was used across the two tasks within the same individuals, our findings strongly suggest that
7 the rule by which pairings involving “less effective” inputs trigger stronger multisensory processing
8 extends beyond immediate behavioural and perceptual benefits, and can also strongly impact future
9 perception and behaviour. Some research investigated if benefits of multisensory memory extend to
10 touch (Lehmann & Murray, 2005) and smell (Gottfried, Smith, Rugg, & Dolan, 2004), but more
11 systematic research is required to draw strong conclusions about the generalisability of the benefits
12 of multisensory memories across the different sensory systems. Likewise, it will be essential to
13 equate baseline performance on tasks in all sensory modalities before more fully invoking inverse
14 effectiveness as an explanation for larger memory benefits on the auditory than visual task in our
15 studies.

2. Brain correlates of implicit multisensory benefits in memory

37
38
39
40
41
42 The majority of our brain mapping studies has focused on the networks involved in visual
43 memory but all our studies employed the continuous old/new recognition paradigm described
44 above (see [Figure 2b](#)). In this section, we focus exclusively on brain responses elicited by *repeated*
45 object presentations. Across both ERP and fMRI methods, portions of the lateral occipital cortex
46 (LOC) were found to respond more strongly to naturalistic visual objects that had been initially
47 accompanied by *semantically* congruent sounds (when compared to repeated imaged that were
48 never presented together with sounds). More specifically, in the ERP study (Murray et al. 2004) we
49 established that distinguishable brain networks (viz. ERP topographic differences) become active
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 already within the first 60–135ms post-stimulus (subsequent effects were also observed at ~210-
2 260ms and 318-390ms). Source estimations identified the LOC as responding significantly more
3
4 strongly towards visual objects previously presented in a semantically congruent multisensory
5
6 context (see [Figure 3](#)).
7

8
9 Subsequently, we have compared memory for objects presented only visually with that for
10
11 objects presented initially in a *meaningless* multisensory context. That is, visual objects were all
12
13 paired with the same, single tone (Lehmann & Murray 2005) or each visual object was paired with a
14
15 distinct tone (with tones modulated in their spectral composition, amplitude envelope and
16
17 waveform type; Thelen, Cappe, & Murray, 2012). As highlighted above, such situations led to
18
19 significant memory impairments when measured across the group of studied participants (Lehman &
20
21 Murray, 2005; Thelen et al., 2012). In Thelen et al. (2012), ERP differences between objects initially
22
23 presented exclusively visually and those presented in a multisensory manner (here with meaningless
24
25 tones) began at ~100ms post-stimulus. As in Murray et al. (2004), these effects were driven by
26
27 changes in the ERP topography, rather than the strength of activation of the electric field at the
28
29 scalp, suggesting that changes in the underlying sources were responsible for distinct memory
30
31 performance. Source estimations of these differences were localised to a small cluster within the
32
33 right LOC (as in Murray et al., 2004) and a larger cluster in the right posterior superior temporal
34
35 sulcus (pSTS). However, in Thelen et al. (2012) the LOC brain activity was weaker in the had-been
36
37 multisensory when compared with the had-been unisensory condition, contrasting with the earlier
38
39 study of Murray et al., (2004) where multisensory pairings had been semantically congruent.
40
41 Responses in the pSTS were significantly stronger for the had-been multisensory than had-been
42
43 unisensory condition. The brain distinguished between these two presentation types again at later
44
45 stages (270–310ms), with differences visible in the ERP topography and with source estimates in this
46
47 time-period localised within the right middle temporal cortex. Notably, the strength of the activity
48
49 within the right middle temporal cortex over the 270-310ms period was directly related to the
50
51 magnitude of performance impairment ($r_{(10)}=0.627$; $p=0.029$; cf. Figure 3 in Thelen et al., 2012).
52
53
54
55
56
57
58
59
60
61

1 Together, these findings suggest that the set of brain areas activated during the visual memory task
2 is modulated by the effectiveness of a given multisensory encoding context on unisensory memory
3 performance, rather than by the simple presence of the preceding multisensory context.
4
5

6
7 The benefits of multisensory processing for (episodic) memory for auditory objects seem to be
8 supported by a relatively different set of brain areas and possibly altogether different brain
9 mechanisms. In Matusz et al. (2015), our participants discriminated naturalistic sounds that could be
10 accompanied by semantically congruent images or scrambled versions of these images and abstract
11 figures. Notably, as in our purely behavioural study (Thelen et al. 2015), the multisensory benefits
12 elicited in auditory memory by the initial presentation of objects in a semantically congruent
13 multisensory manner were stronger than those for visual memory. We found that, yet again, the
14 brain distinguished between objects based on their initial context at very early time points following
15 stimulus presentation (i.e., 35-85 ms post-stimulus; see [Figure 3](#)). However, the effect of initial
16 context on auditory memory involved the right superior temporal cortex (rSTC), the right
17 intraparietal cortex (rIPC), the right inferior occipital cortex, and left frontal cortex. Critically, the
18 superior temporal and intraparietal cortices were those areas modulating in a manner that mapped
19 on to changes in memory performance. Notably, the direction of these modulations was opposite to
20 those found in the visual task in Murray et al. (2004). That is, the activity within these two brain
21 areas was suppressed most strongly for multisensory congruent pairings compared to either
22 auditory-only stimuli or multisensory meaningless pairings. This effect resembled the “response
23 suppression” mechanism (Bergerbest, Ghahremani, & Gabrieli, 2004; M. M. Murray, Camen, Spierer,
24 & Clarke, 2008) frequently reported as underlying short-lived learning effects within the auditory
25 cortices. In Section 5, we discuss the implications of these findings for furthering our understanding
26 of how objects are represented and accessed.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

57 **3. Individual differences in who benefits from multisensory contexts**

58
59
60
61
62
63
64
65

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

22 An important related domain of active inquiry by our group focuses on the hypothesis that an
23 individual's capacity to integrate multisensory information, such as during a simple detection task,
24 may directly scale to how an individual makes use of multisensory experiences to facilitate object
25 recognition and memory. In other words, is one's ability to benefit from multisensory contexts
26 during a memory task based on a more general capacity to integrate multisensory signals, such as

1 simple beeps and flashes presented simultaneously at the same location? One limitation of our prior
2 work is that all data came from different components of the same paradigm (i.e. initial vs. repeated
3 exposures during a continuous recognition task). Thus, it is unknown if links between multisensory
4 processes persist when measured using two or more distinct tasks (each with their own stimulus set,
5 goals, and attentional demands). Also, we do not know the extent to which any links in multisensory
6 integrative capacity manifest specifically at behavioural and/or brain levels. Thus far, we have
7 demonstrated a link between brain activity at one point in time and behaviour at a subsequent time
8 point on the same task. Initial findings indeed point to links between behaviour on a simple
9 detection task and a standardized questionnaire indexing memory function (the mini mental state
10 examination; Eardley et al., submitted).

25 **4. Cognitive mechanisms by which multisensory contexts improve memory**

26
27
28
29
30
31 Before we draw more general conclusions from our findings, we have to note that our
32 paradigm investigates a very specific but ethologically relevant situation. Namely, the task we utilise
33 focuses on episodic memory (*Have you seen this object before in this experimental block?*), and the
34 effectiveness of the processes underlying this memory system is investigated as a function of
35 multisensory processes that are triggered likely outside of the observer's attentional focus (at least
36 in many of our participants). Additionally, we have studied these processes predominantly with
37 naturalistic objects (sounds and schematic drawings), thus likely triggering associations based on
38 long-term, semantic associations between visual and auditory attributes of real-world objects, such
39 as animals or tools (Beauchamp, Argall, Bodurka, Duyn, & Martin, 2004; Laurienti, Kraft, Maldjian,
40 Burdette, & Wallace, 2004; Mahon & Caramazza, 2011; Werner & Noppeney, 2010). The nature of
41 the processes engaged in our paradigm needs to be borne in mind when interpreting the observed
42 brain and behavioural results and placing and interpreting them within the wider background of
43 other studies on learning and memory in multisensory environments. This clarification helps to
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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 & 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 & Giraud, 2006), albeit differing in its local pattern of activity (e.g., Gibson & 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, Moss, Stamatakis, & Tyler, 2006; see also Murray & 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 subserves (some forms of) multisensory memory. However, some of the discrepancies across studies may be explained by contributions of explicit

1 attention to both sensory modalities as well as to the demands of the task (or not) for effortful
2 encoding of the constituents of the multisensory pairings. In this regard, paradigms such as these
3
4 may be building or accessing much richer representations than those at play in a unisensory
5
6 continuous recognition paradigm. The latter instead promotes access to multisensory
7
8 representations that is implicit, which is corroborated by the early and local nature of the
9
10 modulations observed by us within task-relevant sensory cortices.
11
12

13
14 Specifically, our primary finding across both EEG and fMRI studies was that responses to
15
16 repeated presentations of unisensory visual or auditory stimuli were affected implicitly and at early
17
18 latencies by whether or not these stimuli had been previously presented in conjunction with a sound
19
20 or image. This suggests that brain networks responsible for the processing of unisensory stimuli have
21
22 access to multisensory memory representations early on in sensory–cognitive processing.
23
24 Furthermore, our source estimations indicate that this access initially manifests within unisensory
25
26 object recognition areas (as well as IPC in the case of the auditory memory task). We propose that
27
28 this early modulation reflects the rapid reactivation of distinct multisensory and unisensory
29
30 perceptual traces established during initial stimulus presentation. This notion is supported by
31
32 findings from studies of repetition priming both in the visual (e.g., Doniger et al., 2001) as well as
33
34 auditory modality (e.g. De Lucia et al., 2010; Murray et al., 2008), though we have no reason to
35
36 suspect that repetition priming alone could account for our effects (cf. Murray et al., 2004 for a
37
38 more extensive discussion). This proposal of distinct multisensory and unisensory perceptual traces
39
40 is reinforced by two pieces of evidence: (1) that unisensory object recognition areas demonstrate
41
42 auditory–visual convergence, and (2) that multisensory memory representations are both localized
43
44 and distinguishable from their unisensory counterparts. However, and despite evidence from single-
45
46 unit intracranial recordings in non-human primates showing distinct representations for
47
48 multisensory vs. unisensory paired associations (e.g. Gibson and Maunsell, 1997), we cannot fully
49
50 discount the possibility (which may not be mutually exclusive with the above) that the initial
51
52 multisensory experiences are instead impacting unisensory representations.
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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, Retsa, & Murray, 2016; Sarmiento, Matusz, Sanabria, & Murray, 2015; reviewed in Doehrmann & 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 & Gross, 1994; Gibson & Maunsell, 1997; Haenny, Maunsell, & Schiller, 1988; Maunsell, Sclar, Nealey, & DePriest, 1991; see also Goulet & 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.

1 the absence of other activations may also be linked to the fact that their imaging results were
2 masked by the results of the contrast between encoding visual and audio-visual conditions. Findings
3 from these studies were taken as support for the “redintegration” hypothesis (Hamilton, 1859),
4 according to which the repeated presentation of stimuli in the task-relevant sense reactivates both
5 the sensory-specific cortices (here, visual) as well as the secondary-modality cortices (here, auditory)
6 because of a consolidated memory for the audio-visual association. However, it cannot be
7 discounted that the additional activations in these studies were driven by participants using a form
8 of auditory “imagery”, given they had to explicitly recall if the word had been originally learned with
9 a sound. It should also be noted that performance was worse for visually-presented words
10 previously learned with sounds.
11
12
13
14
15
16
17
18
19
20
21
22

23 The fact that our task does not require either effortful study of the multisensory pairs, or
24 attention to the multisensory (or congruent) nature of the pairing, supports the interpretation of our
25 results in terms of multisensory processes involuntarily creating (or accessing) distinct, durable
26 representations for naturalistic images (as well as sounds) when accompanied by (in particular
27 semantically congruent) stimuli in another sense. Direct comparison between our studies and this
28 prior work is difficult based on the blocked nature of the study design and the poor temporal
29 resolution of the prior work. Thus, redintegration processes may play a role in how memories are
30 built and retrieved in multisensory settings. However, our results would necessarily constrain the
31 nature of brain mechanisms governing these processes. First, redintegration may occur via
32 multisensory processing, rather than via concomitant activity between the task-relevant sensory-
33 specific auditory and visual cortices. Even then, our results constrain the possible time course of
34 such effects. That is, if we look at the activations within the auditory cortex reported by Thelen et al.
35 (2012), these were preceded by differential activity within generators localised within the LOC. Thus,
36 any sensory-specific cortex activation would be expected to follow from reactivation of multisensory
37 representations, rather than the initial auditory sensory experience. In other words, the repeated
38 experience of the association may be a consequence, rather than a cause, of successful memory
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 retrieval (see Rugg, Johnson, Park, & Uncapher, 2008 for a similar critique of “reinstatement”
2 accounts of memory retrieval brain mechanisms in visual memory research).
3

4 It is important to consider an alternative interpretation for the basis of these memory
5 benefits, which is that multisensory exposures do not benefit memory processes per se, but rather
6 only perception processes. Any situation, whether a multisensory exposure or some other emphasis
7 of the task-relevant unisensory experience, would be sufficient to elicit the observed positive effects
8 on later recognition memory, because the initial exposure is made more memorable in terms of its
9 perception. While we cannot unequivocally exclude such an account at this stage, several aspects of
10 our data would speak against a purely perceptual account.
11
12
13
14
15
16
17
18
19
20

21 First, on initial stimulus presentations there is no evidence for accuracy differences between
22 unisensory and multisensory conditions (either semantically congruent, incongruent, or entailing a
23 meaningless task-irrelevant stimulus). By contrast, reaction times for all multisensory stimuli were
24 significantly slower than those to unisensory stimuli. This pattern was consistently observed both
25 when the continuous recognition task was performed in the visual modality as well as when the task
26 was performed in the auditory modality (cf. Figure 2 in Thelen and Murray, 2013). If perceptual
27 processes were mediating our effects, then a strong prediction would have been one of more
28 accurate and faster behaviour to multisensory than unisensory stimuli during this initial encounter.
29 Moreover, the similar pattern across all multisensory conditions on initial stimulus presentations
30 does not explain the differential pattern on repeated presentations according to the semantic
31 congruence of the initial multisensory exposure. Second, our ERP effects were consistent in terms of
32 topographic modulations. This was the case both when the task was visual (Murray et al., 2004;
33 Thelen et al., 2012) and when it was auditory (Matusz et al., 2015). Because topographic
34 modulations are forcibly the result of changes in the configuration of underlying brain sources, an
35 explanation based on increased salience or attention (i.e. typical perceptual processes) is unlikely as
36 these have reliably led to changes in ERP amplitude or strength rather than topography. Finally, in
37 addition to investigating ERP effects on stimulus repetitions, we have also reported how ERP
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 responses during initial stimulus presentation are predictive of whether or not an individual's
2 memory is enhanced upon repeated stimulus presentation (Thelen et al., 2014). Individuals who
3
4 show memory enhancements also show stronger responses to initial multisensory stimulus
5 presentations. This was not the case for unisensory stimulus presentations. An account where
6
7 perceptual processes are the root of the enhanced memory would have predicted generally stronger
8
9 responses (both multisensory and unisensory) in individuals exhibiting memory performance
10 enhancement vs. those individual exhibiting memory performance decrements. This was not the
11
12 case.
13
14
15
16
17
18
19
20

21 **5. Broader implications**

22
23 The demonstration of benefits from multisensory contexts on memory advances our
24 understanding of both multisensory processes in general as well as of memory and the organisation
25 of semantic knowledge.
26
27
28
29

30 First, in terms of implications for multisensory processing, the reviewed findings demonstrate
31 that the products of multisensory processes persist over time, influencing subsequent unisensory
32 object perception. Multisensory processes associated with the initial encounter of an object will
33 influence the later retrieval of that object in an involuntary, incidental and general fashion. This
34 underlines the importance of the individual's experience, both long-term as well as short-term (e.g.,
35 inter-trial effects), in influencing responses to both unisensory (visual, auditory, etc.) and
36 multisensory objects. In agreement with these findings, recent theoretical frameworks have aimed
37 to clarify the role of memory and its interplay with other top-down processes, such as selective
38 attention, in controlling distinct multisensory processes (Murray, Lewkowicz, Amedi, & Wallace,
39 2016; ten Oever et al., 2016; see also Matusz, Traczyk, Sobkow, & Strelau, 2015 for evidence for
40 impact of biases in selective attention towards threat-related visual stimuli on their subsequent
41 memory). Understanding the dependencies of different multisensory processes on experience and
42 attention has clear implications for supporting healthy learning as well as rehabilitation of sensory
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 and learning disorders (Bach, Richardson, Brandeis, Martin, & Brem, 2013; Murray, Matusz, &
2 Amedi, 2015).
3

4 Second, these findings highlight the fact that the existing models of memory may not
5 generalise to multisensory settings, particularly when notions such as conceptual novelty vs. physical
6 familiarity come into play (e.g. Reggev et al., 2016). For one, the results reviewed here show that
7 despite their task-irrelevance, semantically congruent contexts bring benefits over and above those
8 predicted by some of the most fundamental hypotheses developed within visual memory research,
9 such as the encoding-retrieval congruence hypothesis (Baddeley et al., 2009; Smith & Vela, 2001). As
10 such, these results bridge traditionally separate lines of research by demonstrating that memory
11 processes may be coupled with more general multisensory processes, such as those readily studied
12 in simple, detection-like perceptual tasks. Further research will be required to ascertain the extent
13 to which the capacity to benefit from multisensory contexts for memory functions is yoked upon a
14 more general capability to benefit from multisensory information in the environment. Ongoing
15 efforts from our group are investigating the extent to which both school-aged children as well as the
16 elderly exhibit such links.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

35 Furthermore, our findings provide direct evidence to the models of functional brain
36 organisation that propose that object representations are quintessentially multisensory. In
37 agreement, recent conceptual frameworks suggest that higher-order visual cortices may be better
38 defined by the tasks they subserve than by the sensory inputs they are traditionally thought to
39 receive (Murray, Lewkowicz, et al., 2016; Reich, Maidenbaum, & Amedi, 2012; ten Oever et al.,
40 2016). In other words, some regions may be specialized to perform a given object recognition
41 function (e.g. face, letters, body parts, etc.), irrespective of whether object-relevant information is
42 conveyed by images, sounds, or touch. This has been demonstrated both when environmental
43 sounds of the objects were presented as well as when same objects were re-coded via sensory
44 substitution devices. For example, Mahon et al. (2009) show that the topological pattern observed in
45 the sighted with visual stimuli was the same as that in the congenitally blind when sounds were
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 used. While memory functions were not explicitly taxed in the majority of these studies, they
2 undoubtedly contribute to the establishment and maintenance of these representations (e.g.,
3
4 Amedi, Raz, Pianka, Malach, & Zohary, 2003). In this perspective, our results enrich our
5
6 understanding of how these multisensory representations are accessed even in unisensory contexts.
7
8 We would underscore that sensory modality does indeed play an important role in memory. It does
9
10 so to the extent that it determines the cortices (and likely also the mechanisms) that will support the
11
12 retrieval of multisensory memories. Likewise, these multisensory contexts provide an implicit “tag”
13
14 to experienced events that can facilitate subsequent recognition. We are currently investigating
15
16 what limitations apply to this tagging and by extension their behavioural consequences. Finally, the
17
18 evidence thus far would indicate that semantic, memory-based processes seem to be most effective
19
20 in conferring benefits upon memory, while processes based on low-level factors such as stimulus
21
22 timing appear to exhibit higher degrees of inter-individual variability. Therefore, it will be crucial to
23
24 determine how to optimize these functions at the service of learning and memory.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

References

- 1
2 Amedi, A., Raz, N., Pianka, P., Malach, R., & Zohary, E. (2003). Early “visual” cortex activation
3 correlates with superior verbal memory performance in the blind. *Nature Neuroscience*, *6*(7),
4 758–766. <http://doi.org/10.1038/nn1072>
5
- 6 Bach, S., Richardson, U., Brandeis, D., Martin, E., & Brem, S. (2013). Print-specific multimodal brain
7 activation in kindergarten improves prediction of reading skills in second grade. *Neuroimage*,
8 *82*, 605–615.
9
- 10 Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Psychology Press.
11
- 12 Baum, S.H., Stevenson, R.A., & Wallace, M.T. (2015). Behavioral, perceptual, and neural alterations
13 in sensory and multisensory function in autism spectrum disorder. *Progress in Neurobiology*,
14 *134*, 140-60.
15
- 16 Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004). Unraveling multisensory
17 integration: Patchy organization within human STS multisensory cortex. *Nature Neuroscience*,
18 *7*(11), 1190–1192. <http://doi.org/10.1038/nn1333>
19
- 20 Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015). Delineating the
21 effect of semantic congruency on episodic memory: The role of integration and relatedness.
22 *PLoS ONE*, *10*(2), 1–24. <http://doi.org/10.1371/journal.pone.0115624>
23
- 24 Bergerbest, D., Ghahremani, D. G., & Gabrieli, J. D. E. (2004). Neural correlates of auditory repetition
25 priming: reduced fMRI activation in the auditory cortex. *Journal of Cognitive Neuroscience*,
26 *16*(6), 966–977.
27
- 28 Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual
29 recognition memory. *Proceedings of the National Academy of Sciences*, *106*(14), 6008–6010.
30 <http://doi.org/10.1073/pnas.0811884106>
31
- 32 Colombo, M., & Gross, C. G. (1994). Responses of inferior temporal cortex and hippocampal neurons
33 during delayed matching-to-sample in monkeys (*Macaca fascicularis*). *Behavioral Neuroscience*,
34 *108*, 443–455.
35
- 36 Constantinescu, A. O., O’Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in
37 humans with a gridlike code. *Science*, *352*(6292), 1464–1468.
38
- 39 Crosse, M. J., Di Liberto, G. M., & Lalor, E. C. (2016). Eye can hear clearly now: inverse effectiveness
40 in natural audiovisual speech processing relies on long-term crossmodal temporal integration.
41 *Journal of Neuroscience*, *36*(38), 9888–9895.
42
- 43 De Lucia, M., Cocchi, L., Martuzzi, R., Meuli, R. A., Clarke, S., & Murray, M. M. (2010). Perceptual and
44 semantic contributions to repetition priming of environmental sounds. *Cerebral Cortex*, *20*(7),
45 1676–1684. <http://doi.org/10.1093/cercor/bhp230>
46
- 47 De Meo, R., Murray, M. M., Clarke, S., & Matusz, P. J. (2015). Top-down control and early
48 multisensory processes: Chicken vs. egg. *Frontiers in Integrative Neuroscience*, *9*, 1–6.
49 <http://doi.org/10.3389/fnint.2015.00017>
50
- 51 Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning
52 modulates processes of audio-visual integration. *Brain Research*, *1242*, 136–150.
53 <http://doi.org/10.1016/j.brainres.2008.03.071>
54
- 55 Doniger, G. M., Foxe, J. J., Schroeder, C. E., Murray, M. M., Higgins, B. A., & Javitt, D. C. (2001). Visual
56 Perceptual Learning in Human Object Recognition Areas: A Repetition Priming Study Using
57
58
59
60
61
62
63
64
65

High-Density Electrical Mapping. *NeuroImage*, 13(2), 305–313.
<http://doi.org/10.1006/nimg.2000.0684>

- 1
2
3 Eardley, A.F., Edgington, T., Oyekan, R., Smyth, E., Murray, M.M., & Matusz, P.J. (2017). Linking
4 multisensory and memory functions in ageing. *Neurobiology of Aging*, in review.
5
6 Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working
7 memory. *Trends in Cognitive Sciences*. <http://doi.org/10.1016/j.tics.2011.11.014>
8
9 Gibson, J.R., & Maunsell, J. H. R. (1997). Sensory Modality Specificity of Neural Activity Related to
10 Memory in Visual Cortex. *Journal of Neurophysiology*, 78, 1263–1275.
11
12 Gingras, G., Rowland, B.A., & Stien, B.E. (2009). The differing impact of multisensory and unisensory
13 integration on behavior. *Journal of Neuroscience* 29: 4897-4902.
14
15 Gottfried, J. A., Smith, A. P. R., Rugg, M. D., & Dolan, R. J. (2004). Remembrance of odors past:
16 Human olfactory cortex in cross-modal recognition memory. *Neuron*, 42(4), 687–695.
17 [http://doi.org/10.1016/S0896-6273\(04\)00270-3](http://doi.org/10.1016/S0896-6273(04)00270-3)
18
19 Goulet, S., & Murray, E. A. (2001). Neural substrates of crossmodal association memory in monkeys:
20 the amygdala versus the anterior rhinal cortex. *Behavioral Neuroscience*, 115(2), 271.
21
22 Haenny, P. E., Maunsell, J. H. R., & Schiller, P. H. (1988). State dependent activity in monkey visual
23 cortex II. Retinal and extraretinal factors in V4. *Experimental Brain Research*, 69, 245–259.
24
25 Hamilton, W. (1859). *Lectures on metaphysics and logic* (Vol. 1). Boston, MA (USA): Gould and
26 Lincoln.
27
28 Heikkilä, J., Alho, K., Hyvönen, H., & Tiippana, K. (2015). Audiovisual semantic congruency during
29 encoding enhances memory performance. *Experimental Psychology*, 62(2), 123–130.
30 <http://doi.org/10.1027/1618-3169/a000279>
31
32 Heikkilä, J., & Tiippana, K. (2016). School-aged children can benefit from audiovisual semantic
33 congruency during memory encoding. *Experimental Brain Research*, 234(5), 1199–1207.
34 <http://doi.org/10.1007/s00221-015-4341-6>
35
36 Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W., & Stein, B. E. (2001). Two cortical areas mediate
37 multisensory integration in superior colliculus neurons. *Journal of Neurophysiology*, 85, 506–
38 522.
39
40 Johansson, B.B. (2012). Multisensory stimulation in stroke rehabilitation. *Front Hum Neurosci*, 6, 60.
41
42 Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic
43 congruence is a critical factor in multisensory behavioral performance. *Experimental Brain*
44 *Research*, 158(4), 405–414. <http://doi.org/10.1007/s00221-004-1913-2>
45
46 Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object
47 discrimination. *Cognitive Brain Research*, 24(2), 326–334.
48 <http://doi.org/10.1016/j.cogbrainres.2005.02.005>
49
50 Lewis, J.W. (2010). Audio-visual perception of everyday natural object - hemodynamic studies in
51 humans. In: *Multisensory Object Perception in the Primate Brain*, Naumer and Kaiser (Eds.).
52 New York (USA): Springer.
53
54 Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.
55
56 Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the
57 brain? The distributed domain-specific hypothesis. *Trends in Cognitive Sciences*, 15, 97–103.
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

<http://doi.org/10.1016/j.tics.2011.01.004>.What

- 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. *European Journal of Neuroscience*. <http://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. *Journal of Cognitive Psychology*. <http://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. *Visual Neuroscience*, *7*(6), 561–573. <http://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://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. *Multisensory Research*, *26*, 581–592. <http://doi.org/10.1163/22134808-00002436>
- Murray, E. A., & Bussey, T. J. (1999). Perceptual-mnemonic functions of the perirhinal cortex. *Trends in Cognitive Sciences*, *3*(4), 142–151. [http://doi.org/10.1016/S1364-6613\(99\)01303-0](http://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., & Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage*, *21*(1), 125–135. <http://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://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*. Boca Raton, FL (USA): CRC Press.
- Murray, M. M., Matusz, P. J., & Amedi, A. (2015). Neuroplasticity: Unexpected Consequences of Early Blindness. *Current Biology*, *25*(20), R998–R1001. <http://doi.org/10.1016/j.cub.2015.08.054>
- Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2016a). 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. (2016b). Multisensory Processes: A Balancing Act across the Lifespan. *Trends in Neurosciences*, *39*(8), 567–579. <http://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. *Neuroscience Letters*, *505*(2), 191–195. <http://doi.org/10.1016/j.neulet.2011.10.022>

- 1 Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain
2 activity during memory retrieval. *Proceedings of the National Academy of Sciences*, *97*, 11120–
3 11124. <http://doi.org/10.1073/pnas.97.20.11120>
- 4 Reggev, N., Bein, O., & Maril, A. (2016). Distinct neural suppression and encoding effects for
5 conceptual novelty and familiarity. *Journal of Cognitive Neuroscience*
6 *Journal of Cognitive Neuroscience*, *28*, 1455–1470. <http://doi.org/10.1162/jocn>
7
- 8 Reich, L., Maidenbaum, S., & Amedi, A. (2012). The brain as a flexible task machine: Implications for
9 visual rehabilitation using noninvasive vs. invasive approaches. *Current Opinion in Neurology*,
10 *25*(1), 86–95. <http://doi.org/10.1097/WCO.0b013e32834ed723>
- 11 Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2009). Preperceptual and Stimulus-Selective
12 Enhancement of Low-Level Human Visual Cortex Excitability by Sounds. *Current Biology*, *19*(21),
13 1799–1805. <http://doi.org/10.1016/j.cub.2009.09.027>
- 14 Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2013). The Contributions of Sensory Dominance and
15 Attentional Bias to Cross-modal Enhancement of Visual Cortex Excitability. *Journal of Cognitive
16 Neuroscience*, *25*, 1122–1135. <http://doi.org/10.1162/jocn>
- 17 Rugg, M. D., Johnson, J. D., Park, H., & Uncapher, M. R. (2008). Encoding-retrieval overlap in human
18 episodic memory: A functional neuroimaging perspective. *Progress in Brain Research*, *169*,
19 339–352. [http://doi.org/10.1016/S0079-6123\(07\)00021-0](http://doi.org/10.1016/S0079-6123(07)00021-0)
- 20 Sarmiento, B. R., Matusz, P. J., Sanabria, D., & Murray, M. M. (2015). Contextual factors multiplex to
21 control multisensory processes. *Human Brain Mapping*, *37*, 273–288.
22 <http://doi.org/10.1002/hbm.23030>
- 23 Schmid, C., Büchel, C., & Rose, M. (2011). The neural basis of visual dominance in the context of
24 audio-visual object processing. *NeuroImage*, *55*(1), 304–11.
25 <http://doi.org/10.1016/j.neuroimage.2010.11.051>
- 26 Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory. *Psychonomic Bulletin &
27 Review*, *8*, 203–220. <http://doi.org/10.3758/BF03196157>
- 28 Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., ... Wallace, M. T.
29 (2014). Identifying and Quantifying Multisensory Integration: A Tutorial Review. *Brain
30 Topography*, *27*(6), 707–730. <http://doi.org/10.1007/s10548-014-0365-7>
- 31 Stevenson, R. A., & James, T. W. (2009). Audiovisual integration in human superior temporal sulcus:
32 Inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage*,
33 *44*(3), 1210–1223. <http://doi.org/10.1016/j.neuroimage.2008.09.034>
- 34 Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and
35 computational mechanisms. *Nature Reviews Neuroscience*, *(15)*, 745–756.
36 <http://doi.org/10.1038/nrn3838>
- 37 Taylor, K. I., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2006). Binding crossmodal object features in
38 perirhinal cortex. *Proceedings of the National Academy of Sciences of the United States of
39 America*, *103*(21), 8239–8244. <http://doi.org/10.1073/pnas.0509704103>
- 40 ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M. M., & Matusz, P. J. (2016).
41 The COGs (context, object, goals) in multisensory processing. *Experimental Brain Research*, *234*,
42 1307–1323. <http://doi.org/10.1007/s00221-016-4590-z>
- 43 Thelen, A., Cappe, C., & Murray, M. M. (2012). Electrical neuroimaging of memory discrimination
44 based on single-trial multisensory learning. *NeuroImage*, *62*, 1478–1488.

<http://doi.org/10.1016/j.neuroimage.2012.05.027>

- 1
2 Thelen, A., Matusz, P. J., & Murray, M. M. (2014). Multisensory context portends object memory. *Current Biology*, *24*, R434-5. Retrieved from
3 <http://www.sciencedirect.com/science/article/pii/S0960982214007568>
4
5
6 Thelen, A., & Murray, M. M. (2013). The Efficacy of Single-Trial Multisensory Memories. *Multisensory
7 Research*, *26*(5), 483–502. <http://doi.org/10.1163/22134808-00002426>
8
9 Thelen, A., Talsma, D., & Murray, M. M. (2015). Single-trial multisensory memories affect later
10 auditory and visual object discrimination. *Cognition*, *138*, 148–160.
11 <http://doi.org/10.1016/j.cognition.2015.02.003>
12
13 Ueno, D., Masumoto, K., Sutani, K., & Iwaki, S. (2015). Latency of modality-specific reactivation of
14 auditory and visual information during episodic memory retrieval. *Neuroreport*, *26*(6), 303–
15 308. <http://doi.org/10.1097/WNR.0000000000000325>
16
17 Von Kriegstein, K., & Giraud, A. L. (2006). Implicit multisensory associations influence voice
18 recognition. *PLoS Biology*, *4*(10), 1809–1820. <http://doi.org/10.1371/journal.pbio.0040326>
19
20 Wallace, M. T. (2004). The development of multisensory processes. *Cognitive Processing*, *5*(2), 69–
21 83. <http://doi.org/10.1007/s10339-004-0017-z>
22
23 Wallace, M. T., Carriere, B. N., Perrault, T. J., Vaughan, J. W., & Stein, B. E. (2006). The development
24 of cortical multisensory integration. *The Journal of Neuroscience*, *26*(46), 11844–11849.
25 <http://doi.org/10.1523/JNEUROSCI.3295-06.2006>
26
27 Wallace, M. T., Meredith, M. A., & Stein, B. E. (1998). Multisensory integration in the superior
28 colliculus of the alert cat. *Journal of Neurophysiology*, *80*, 1006–1010.
29
30 Wallace, M. T., Perrault, T. J., Hairston, W. D., & Stein, B. E. (2004). Visual experience is necessary for
31 the development of multisensory integration. *The Journal of Neuroscience : The Official Journal
32 of the Society for Neuroscience*, *24*(43), 9580–9584. [http://doi.org/10.1523/JNEUROSCI.2535-
33 04.2004](http://doi.org/10.1523/JNEUROSCI.2535-04.2004)
34
35
36 Werner, S., & Noppeney, U. (2010). Distinct functional contributions of primary sensory and
37 association areas to audiovisual integration in object categorization. *The Journal of
38 Neuroscience*, *30*(7), 2662–2675. <http://doi.org/10.1523/JNEUROSCI.5091-09.2010>
39
40 Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering
41 reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*,
42 11125–11129.
43
44 Yuval-Greenberg, S., & Deouell, L. Y. (2007). What You See Is Not (Always) What You Hear: Induced
45 Gamma Band Responses Reflect Cross-Modal Interactions in Familiar Object Recognition. *The
46 Journal of Neuroscience*, *27*(5), 1090–1096. <http://doi.org/10.1523/jneurosci.4828-06.2007>
47
48 Yuval-Greenberg, S., & Deouell, L. Y. (2009). The dog's meow asymmetrical interaction in cross-
49 modal object recognition. *Experimental Brain Research*, *193*(4), 603–614.
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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.

Figure & table captions

Figure 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.

Figure 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.

Figure 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 and Matusz et al., 2015).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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 & Murray (2005); Naghavi et al. (2011); Moran et al. (2013); Thelen et al. (2012, 2015); Thelen, Matusz, Murray (2014); Matusz, Thelen et al. (2015); Heikkilä et al. (2015, 2016)
Continuous recognition	Study and test items are intermixed.	✓	Murray et al. (2004, 2005); Lehmann & Murray (2005); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen, Matusz, Murray (2014); Matusz, Thelen et al. (2015)
Single-trial	Initial multisensory presentations occur once.	✓	Murray et al. (2004, 2005); Lehmann & Murray (2005); Naghavi et al. (2011); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen, Matusz, Murray (2014); Matusz, Thelen et al. (2015)
Semantic congruence	The initial presentation consists of stimuli referring to the same object.	✓	Murray et al. (2004, 2005); Lehmann & Murray (2005; Exp.2); Naghavi et al. (2011); von Kriegstein et al. (2006); Moran et al. (2013); Thelen et al. (2015); Matusz, Thelen et al. (2015); 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 & Murray (2005); von Kriegstein et al. (2006); Naghavi et al. (2011); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen, Matusz, Murray (2014); Matusz, Thelen et al. (2015); Ueno et al. (2015); Heikkilä et al. (2015, 2016)
Variation in multisensory pairings	Stimuli in the task-irrelevant sense vary across trials	✗†	Murray et al. (2004, 2005); Gottfried et al. (2004); Lehmann & Murray (2005; Exp.2); von Kriegstein et al. (2006); Naghavi et al. (2011); Thelen et al. (2012, 2015); Moran et al. (2013); Thelen, Matusz, Murray (2014); Matusz, Thelen et al. (2015); 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 minute	Murray et al. (2004, 2005); Lehmann & Murray (2005; Exp.2); Thelen et al. (2015); Moran et al. (2013); Thelen, Matusz, Murray (2014); Matusz, Thelen et al. (2015)
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 & Murray (2005; Exp.2); Thelen et al. (2015); Moran et al. (2013); Thelen, Matusz, Murray (2014); Matusz, Thelen et al. (2015);
Persistence over multiple sessions	Benefits extend beyond a single testing session (in continuous recognition)	?	-
Physical vs. conceptual initial-repeated semantic congruence	Initial and repeated presentations are physically identical or refer to the same object, parametrically varying semantic relatedness	?	-

†This factor exhibits high inter-individual variability.

Figure 1



Figure 2

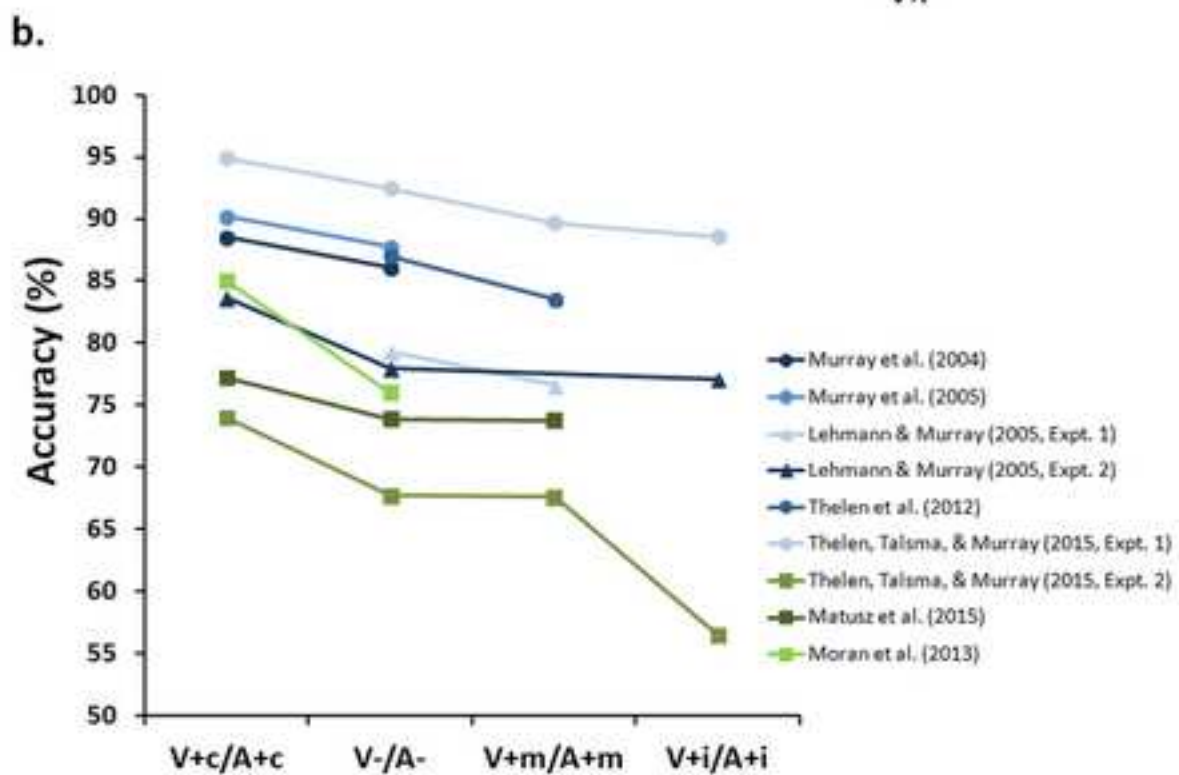
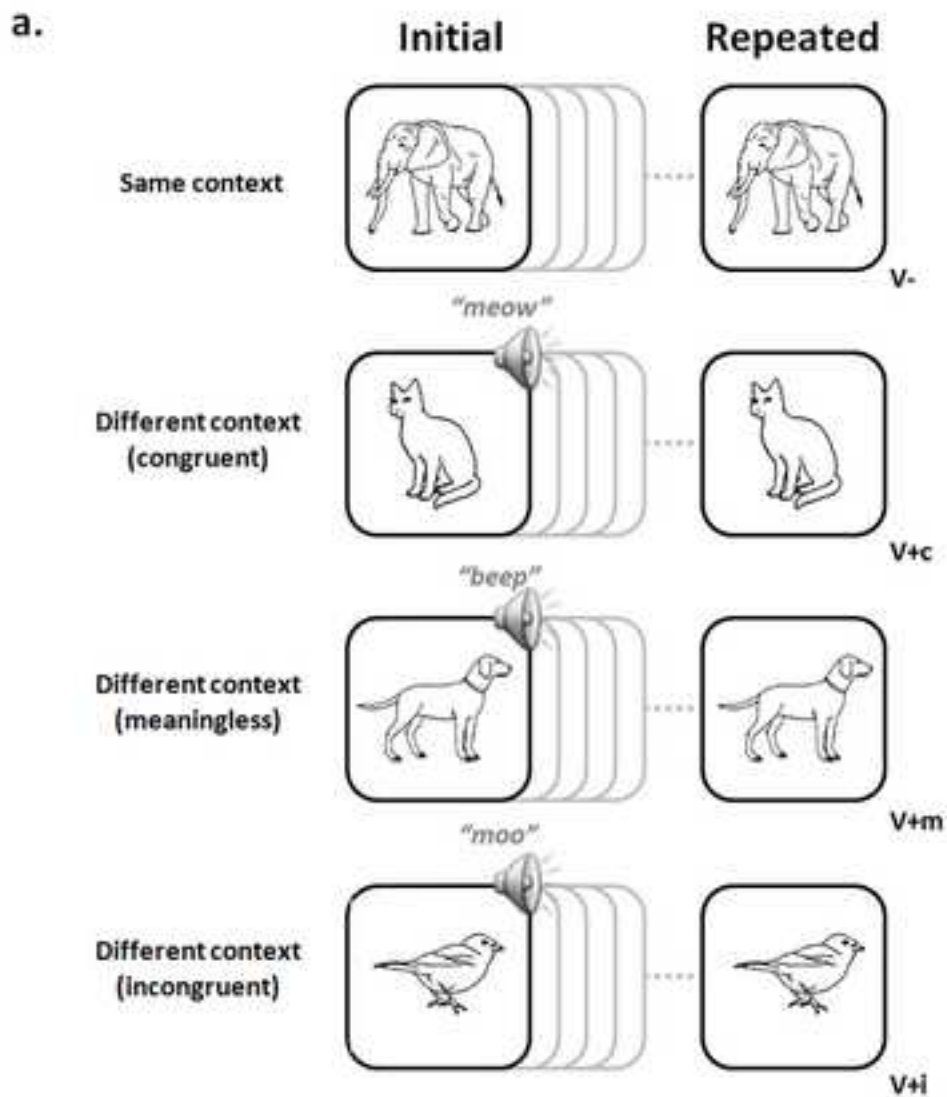


Figure 3

