



Does multisensory study benefit memory for pictures and sounds?

Diane Pecher^{*}, René Zeelenberg

Erasmus University Rotterdam, The Netherlands

ARTICLE INFO

Keywords:

Multisensory memory
Multisensory integration
Encoding specificity
Continuous recognition

ABSTRACT

Studies have found a multisensory memory benefit: higher recognition accuracy for unimodal test items that were studied as bimodal items than for those studied as unimodal items. This is a surprising finding because the encoding specificity principle predicts that memory performance should be better with greater overlap between processing during study and test. We used Thelen, Talsma, and Murray's (2015) method who previously found a multisensory memory benefit. Items were presented as unimodal (picture or sound) or bimodal (picture and sound) items in a continuous recognition task in which only one modality was task-relevant. In four experiments we obtained little evidence for a difference in memory performance between items studied as unimodal or bimodal stimuli, but there was a benefit of study-test overlap in format if sound was the task-relevant modality. Task-induced attention for the irrelevant modality or response bias may have played a role in previous studies. We conclude that the multisensory memory benefit may not be a general finding, but rather one that is found only under conditions that induce participants to pay attention to the task-irrelevant modality.

Memory might be better for events that are experienced in multiple sensory modalities than for events that are experienced in a single sensory modality (Meyerhoff & Huff, 2016; von Kriegstein & Giraud, 2006; see Shams & Seitz, 2008 for a review). The most striking finding is that this benefit of multisensory learning is found even when participants are tested with unimodal targets. For example, when participants are tested with a picture of a dog, they are more likely to recognize the picture as a studied item if earlier on the picture had been presented together with the sound of a dog than if it had been presented without a sound (Heikkilä, Alho, Hyvönen, & Tiippana, 2015; Lehmann & Murray, 2005; Meyerhoff & Huff, 2016; Moran et al., 2013; Thelen et al., 2015).

These findings are in contrast with the idea that memory performance benefits from an overlap in the type of processing between study and test, formulated as the encoding specificity principle (Tulving & Thomson, 1973) and the principle of transfer appropriate processing (Morris, Bransford, & Franks, 1977). These principles are supported by many findings that memory performance is positively related to the overlap in processing between study and test (e.g., Barclay, Bransford, Franks, McCarrel, & Nitsch, 1974; Blaxton, 1989; Cabeza, 1994; Parks, 2013; Pecher, Zeelenberg, & Barsalou, 2004; Pecher, Zanolie, & Zeelenberg, 2007; Roediger & Adelson, 1980; Roediger & Blaxton, 1987; Van Dantzig, Cowell, Zeelenberg, & Pecher, 2011; Zeelenberg, 2005; Zeelenberg, Pecher, Shiffrin, & Raaijmakers, 2003; see also Roediger,

Weldon, & Challis, 1989). One interpretation of these findings is the context in which an item is presented and the task instructions affect what features are attended and encoded in memory and consequently determine what is an effective retrieval cue during test. Following these principles, we would expect better memory performance for stimuli with greater overlap between study and test. Participants are likely to process a bimodal stimulus differently than a unimodal stimulus, for example they might attend different features if a picture is presented with a sound than if it is presented without a sound. Several studies have shown that memory for words is affected by the overlap in the modality specific meaning that was activated during study and test (Pecher et al., 2004; Van Dantzig et al., 2011). During test, when the picture is presented without sound, the features activated by the test item will be more similar to those activated during unimodal study than during bimodal study. Thus, for items that are tested as unimodal items performance should be better if they are studied as unimodal items than if they are studied as multimodal items. The finding of the opposite result, namely a benefit for multisensory items, deviates from this general memory principle. The finding of a robust multisensory memory benefit that generalizes across stimuli and procedures would be interesting and stimulate further research into the processes driving the effect. Moreover, a robust multisensory memory benefit might be helpful in developing methods that optimize learning and retention of new information

^{*} Corresponding author at: Department of Psychology, Child and Educational Studies, room T13-48, Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands

E-mail address: pecher@essb.eur.nl (D. Pecher).

<https://doi.org/10.1016/j.cognition.2022.105181>

Received 19 October 2021; Received in revised form 16 May 2022; Accepted 23 May 2022

Available online 28 May 2022

0010-0277/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

(e.g., Mayer, 2008).

One possibility is that the benefit for multimodal study items is due to integrative processing that connects the information from different modalities into a single object representation (Quak, London, & Talsma, 2015). Shams and Seitz (2008) have proposed two such integrative mechanisms that might account for the multisensory advantage. Their proposals are based on the idea that information is processed in modality-specific brain areas that are connected through higher level multisensory areas (like convergence zones, Binder & Desai, 2011, Damasio, 1989). Such multisensory areas are activated when information from different sensory modalities is integrated (Woods & Newell, 2004). One proposed explanation for the multisensory advantage is that, during study, activation in the non-tested modality causes stronger activation of the tested modality via these multisensory areas, which results in a stronger memory trace than when information from only one modality is presented. The other proposed explanation is that a multimodal stimulus leads to a stronger memory trace because there is additional encoding in multisensory areas, and during test the unimodal stimulus activates this entire multisensory memory (Shams & Seitz, 2008). The latter explanation is supported by auditory cortex activation during test of words that were studied with a sound compared to words that were studied without a sound (Nyberg, Habib, McIntosh, & Tulving, 2000). Other studies too have shown that stimuli from one modality may activate brain regions associated to a different modality when the items were previously studied as multisensory items including that modality (see Thelen & Murray, 2013, for a review). Although this finding is consistent with the idea that a unisensory stimulus can activate a multisensory memory, such brain activity alone does not show that memory performance is better for multisensory than unisensory items. Both integrative explanations are based on the claim that multisensory integration is essential for the multisensory benefit.

Thelen et al. (2015) further argued that integrative processes will result in stronger memories when information from different modalities is congruent (e.g., a picture of a dog and the sound of a barking dog), that is, activates that same object in long-term memory, but will lead to interference when they are incongruent (e.g., a picture of a dog and the sound of a train). Thus, memory is expected to be better for congruent multimodal items than for unimodal items because a stronger representation is encoded for the target, but worse for incongruent multimodal items than for unimodal items because there is interference from the incongruent object.

Although there is some evidence for a benefit of multisensory learning (Lehmann & Murray, 2005; Matusz et al., 2015; Meyerhoff & Huff, 2016; Moran et al., 2013; Thelen et al., 2015), not all studies have obtained such a benefit (Cohen, Horowitz, & Wolfe, 2009; Nyberg et al., 2000). Moreover, although Thelen et al. observed a negative effect of incongruent multimodal items compared to unimodal items, others found no disadvantage for multimodal multisensory items compared to unimodal items (Heikkilä et al., 2015; Heikkilä, Alho, & Tiippana, 2017; Lehmann & Murray, 2005; Matusz et al., 2015; Moran et al., 2013). Canits, Pecher, and Zeelenberg (2018) found no benefit of congruent motor actions on later memory for pictures even though congruency did affect immediate motor responses during the study task. Thus, the currently available evidence from only a few studies partially supports the idea that multisensory integration explains the memory advantage for multimodal items. The advantage for congruent stimuli is consistent with this explanation, but the absence of a disadvantage for incongruent stimuli (in most published studies) might be hard to explain.

In the current study we aimed to further investigate the effect of multimodal stimulus presentation on memory. Although studies have obtained a multisensory benefit, most of these did not use an optimal design to allow strong conclusions. First, it is often not clear if materials were properly counterbalanced. In most studies, the numbers of sound items and picture items do not match such that a fully counterbalanced design would be possible (Lehmann & Murray, 2005; Murray et al., 2004; Murray, John, & Wylie, 2005; Thelen et al., 2015; Heikkilä et al.,

2015, 2017; Heikkilä & Tiippana, 2016). If some items are systematically presented as unimodal and other items are systematically presented as bimodal a confound exists between stimulus materials and experimental condition. It seems possible that items in one condition are more memorable than those in another condition. Second, with the exception of Thelen et al. (2015), in the continuous recognition paradigm presentation format (unimodal vs. bimodal) and study status (studied vs. nonstudied) were correlated because study items could be unimodal or bimodal, but test items were always unimodal (Lehmann & Murray, 2005; Matusz et al., 2015; Moran et al., 2013; Murray et al., 2004, Murray et al., 2005). In a continuous recognition paradigm (Shepard & Teghtsoonian, 1961), as was used in all five aforementioned studies, the participant has to distinguish study (initial presentation) and test (repeated presentation) items mixed in the same continuous stream of items. Because all bimodal items were study items, the correct response to a bimodal item was always 'new'. Because only a portion (varying from 25 to 50%) of the study items were unimodal but all test items were unimodal, the probability that a unimodal item was old was larger than 50% (varying from 67 to 80%). As a result of this correlation participants may have had a bias to respond 'old' to unimodal items. Moran et al. indeed observed a higher false alarm rate for unimodal than for the bimodal items which might indicate such a bias. The effect of such a bias might be that on their first presentation, unimodal and bimodal items are processed differently. For example, participants may put less effort into encoding unimodal study items than bimodal study items because they are seen as test items. Moreover, when all test items are unimodal this causes multimodal items to be less frequent and, therefore, they may be more distinctive than unimodal items. Memory is better for distinctive items (Hunt & McDaniel, 1993) and thus having unequal numbers of unimodal and bimodal items might lead to differences in memory performance that are unrelated to their modality perse. To summarize, many findings of multisensory advantage may still be open to alternative explanations because of suboptimal designs. Therefore, we conclude that the evidence for a multisensory advantage is not very strong.

In the present study, we wanted to collect more data to establish whether, compared to unimodal items, memory is better for bimodal-congruent items and worse for bimodal-incongruent items. We replicated the continuous recognition experiment by Thelen et al. (2015) with some modifications. We chose the study by Thelen et al. because they presented the same number of unimodal and multimodal items during study and test and therefore their design did not suffer from a correlation between presentation format and study status. Thelen et al. (2015) presented object pictures and sounds in a continuous recognition task in which only one modality was task relevant (i.e., participants made recognition decisions to either pictures or sounds). Participants had to decide if the item was presented for the first time (i.e., a study trial) or repeated (i.e., a test trial). Half of the items were presented as unimodal stimuli and the other half as bimodal stimuli where the irrelevant sensory modality was congruent, incongruent, or meaningless. We dropped the meaningless condition so that more items could be presented in the congruent and incongruent conditions. Unlike Thelen et al., who presented items several times in different blocks, we presented each item only once for study and once for test. In addition, unlike Thelen et al., who tested the same participants in the picture memory and sound memory tasks, creating even more repetitions of items, we tested separate groups of participants in the picture memory task (Experiment 1) and the sound memory task (Experiment 2). Thelen et al. obtained an interaction between task order and presentation format on the study trial. Although they did not present all data that would be needed to fully understand the interaction, it appears that the multisensory advantage in picture memory was restricted to participants who had done the auditory memory task first. This suggests that auditory information may have been attended more by participants who had previously been instructed to remember these sounds than by participants who did the picture memory task first. Another important change

was that we removed a potential confound in stimulus materials as described in the previous paragraph. The appendix in Thelen et al. shows that they presented 144 different pictures but only 106 different sounds. Although they did not fully specify how items were counter-balanced and we thus do not know the details, the different numbers of pictures and sounds suggest that 38 pictures were presented only in the visual-only condition and therefore that there was a confound between materials and study condition in their experiment. We created a set of stimuli with equal numbers of pictures and sounds such that each picture was paired to a sound. We used a random selection procedure such that all items were equally likely to be presented in a specific condition. Finally, Thelen et al. analyzed only data for the items that were presented as unimodal items on the repeated (i.e., second) presentation, but we also looked at the items in multimodal conditions to assess the effect of overlap between study and test. We tested at least twice as many participants per experiment as Thelen et al. who tested 26 participants in their experiment. A replication of Thelen et al.'s findings would show an advantage for unimodal test items studied as bimodal-congruent items and a disadvantage for items studied as bimodal-incongruent items compared to items studied as unimodal items. However, based on the literature on transfer appropriate processing we might expect a different pattern of results, namely that the advantage of bimodal study items would be restricted to items that were also tested as bimodal items.

1. Experiment 1

1.1. Method

1.1.1. Participants

Fifty-four students at the Erasmus University participated for course credit.

1.1.2. Stimuli

A set of 232 picture-sound pairs was used as congruent pairs. Of these, 224 were used in the experimental trials and 8 were used for practice. Of these pairs, 168 were from Moran et al. (2013), kindly provided by Zachary Moran. The other 64 pairs were created by pairing pictures and sounds, similar in quality to those provided by Moran, retrieved from various websites. The pictures were colored line drawings of animals (e.g., dog, bee) and objects (e.g., trumpet, helicopter). The sound files were 500 ms clips of a typical sound made by the animal or object. Incongruent picture-sound pairs were created by randomly pairing pictures to sounds with the restriction that the sound could not be that of a closely related item (e.g., a picture of a trumpet with the sound of a saxophone). A full list of items is provided in Appendix A. For each participant, a different random allocation of items to conditions was created.

1.1.3. Design

Each picture was presented twice. In the continuous recognition task, all items could be considered test items, where the response to an initial presentation should be "new" and the response to a repeated item should be "old". We will call the initial presentation of an item a study trial because this is when the item is initially encoded in memory. We will call the repeated presentation of an item a test trial, because responses on these trials indicate if participants recognize the item as being presented previously. It is important to note that study and test trials were mixed into a single sequence (i.e., there are no separate study and test blocks) and the task for the participant was the same on all trials, namely decide whether the item was new or old. For both study and test trials, half of the items were presented alone as a picture, the unimodal condition. The other half was presented with a sound, the bimodal condition. Half of the sounds were congruent with the picture and half were incongruent. Of all the items in each of the three study conditions (unimodal, bimodal-congruent, bimodal-incongruent), half were subsequently presented at test alone as a picture and half were presented

with a sound (congruent or incongruent). The full set of conditions with number of items in each condition is presented in Table 1.

1.1.4. Procedure

The experimental procedure was closely based on that of Thelen et al. (2015). Participants were tested individually. They were seated at a desk wearing headphones with a keyboard and monitor on the desk in front of them, all connected to the PC that was used to run the experiment. They were instructed that a list of pictures that they had to remember would be shown on the monitor, and more specifically, that they had to decide for each picture whether it was 'old' or 'new'. Participants were informed that some of the pictures were presented together with a sound, but that the old-new decision should be based only on the picture. The experiment started with 16 practice trials, followed by two blocks of 224 experimental trials. Each trial started with a 500 ms fixation marker (+) in the center of the screen, followed by a picture that was shown for 500 ms. If a sound was played, it started at the same time as the picture and played for 500 ms. Each picture was followed by a variable ISI of 900–1500 ms (mimicking the timing used by Thelen et al.) until the next trial started. Participants kept their index fingers on the z and m keys throughout the block. They pressed the z key for new items and the m key for old items. Feedback was provided only during the practice trials. After each trial, the word "Correct!" in blue or "Incorrect" in red was displayed in the center of the screen for 250 ms. During the experimental trials no feedback was provided. Items from different conditions were mixed in a semi-random order. Repetitions were presented at a lag that varied from 5 to 15 items.

1.2. Results

Because participants had to respond before the start of the next trial there were trials on which no response was recorded. Only trials on which a participant responded were included in the following calculations and analyses (94.9% of all trials). The accuracy for new and old items was calculated for each condition. Data from three participants were removed from the analyses because their accuracy was below 60% (note that 50% accuracy represents chance performance). All data that were used for the analyses are available on <https://osf.io/vkqbf/>.

The mean hit rates and false alarm rates for the three study conditions are presented in Table 2. For each type of trial d' (d-prime) was calculated using the Snodgrass and Corwin (1988) correction. Because hit rates (and false alarms) may be influenced by response bias, d' provides a better measure of memory sensitivity than just hit rates. The false alarm rates are calculated from 'old' responses to the first presented items grouped by study condition, so to calculate d' the comparison is between first presentation and second presentation of items in the same format. This was the best comparison to eliminate the effect of response bias toward a particular format. The average d' s for all conditions are shown in Fig. 1. A 3 (study condition) by 3 (test condition) ANOVA showed an effect of study condition, $F(2,100) = 4.89, p = .009$, partial $\eta^2 = 0.09$. It seems that memory was slightly better for items that were studied as picture-only than for items studied as a picture with sound, which is opposite from the previously reported multisensory advantage. Memory was better for items tested with sound than for items tested as picture-only, $F(2,100) = 3.63, p = .030$, partial $\eta^2 = 0.07$. There was no interaction between study and test condition, $F(4,200) = 1.13, p = .344$,

Table 1

Presentation conditions and number of items in each test condition for experiments 1–4.

Study	Test		
	Unimodal	Bimodal congruent	Bimodal incongruent
Unimodal	56	28	28
Bimodal congruent	28	14	14
Bimodal incongruent	28	14	14

Table 2
Mean hit rates and false alarm rates in experiments 1 to 4 (standard errors of the mean in parentheses).

	Study	Hit rates			False alarms	
		Test unimodal	Test congruent	Test incongruent		
Experiment 1	Unimodal	0.881 (0.013)	0.910 (0.011)	0.900 (0.011)	0.092 (0.010)	
	Congruent	0.865 (0.016)	0.898 (0.019)	0.870 (0.017)	0.093 (0.014)	
	Incongruent	0.874 (0.012)	0.904 (0.011)	0.908 (0.018)	0.092 (0.014)	
Experiment 2	Unimodal	0.751 (0.018)	0.675 (0.022)	0.708 (0.019)	0.330 (0.016)	
	Congruent	0.762 (0.017)	0.849 (0.016)	0.682 (0.027)	0.295 (0.018)	
	Incongruent	0.745 (0.019)	0.654 (0.027)	0.835 (0.017)	0.304 (0.019)	
Experiment 3	Unimodal	0.937 (0.007)	0.937 (0.008)	0.942 (0.009)	0.051 (0.006)	
	Congruent	0.948 (0.007)	0.957 (0.008)	0.943 (0.012)	0.043 (0.006)	
	Incongruent	0.940 (0.008)	0.932 (0.012)	0.956 (0.009)	0.041 (0.005)	
Experiment 4	Unimodal	0.786 (0.015)	0.718 (0.019)	0.759 (0.017)	0.256 (0.015)	
	Congruent	0.816 (0.014)	0.903 (0.013)	0.766 (0.021)	0.204 (0.013)	
	Incongruent	0.773 (0.016)	0.737 (0.021)	0.846 (0.015)	0.226 (0.014)	

Note. Hit and false alarm rates reflect values obtained before applying the Snodgrass and Corwin (1988) correction.

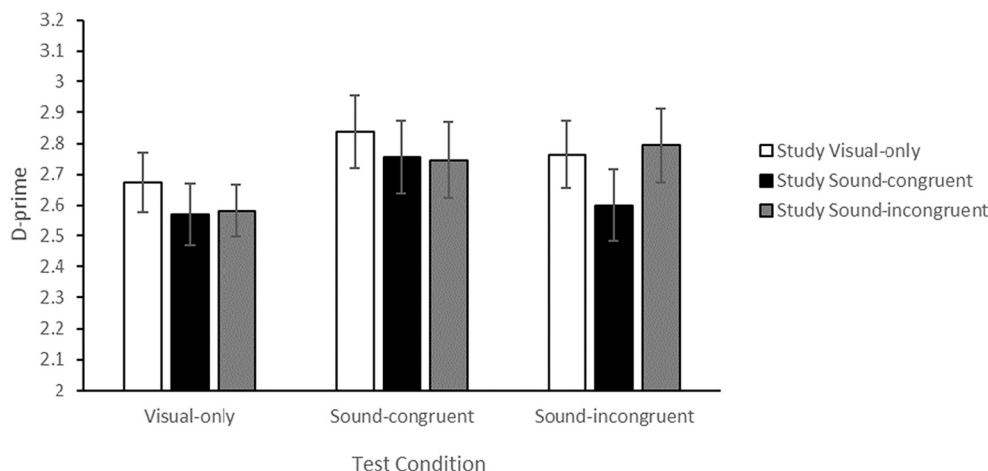


Fig. 1. Average D-primes in the picture recognition memory task of experiment 1 (error bars represent standard errors of the mean).

partial $\eta^2 = 0.02$. The absence of an interaction indicates that the overlap between study and test format did not affect memory.

Because Thelen et al. (2015) analyzed only responses to items tested as unimodal stimuli, we did a second analysis to separately test the effect of study condition on items tested as picture-only. In addition to calculating p values, we calculated the JZS Bayes Factor (BF), which is the ratio of $p(D|H0)$ and $p(D|H1)$, the probabilities of observing the data under the null hypothesis and the alternative hypothesis, respectively. The Bayes Factor thus provides a relative measure of the extent to which the data provide evidence for the null hypothesis of no effect or the alternative hypothesis (Rouder, Speckman, Sun, Morey, & Iverson, 2009). Bayes Factors between 3 and 10 can be considered moderate evidence, and Bayes Factors above 10 can be considered strong evidence. Bayes Factors were calculated using JASP (JASP Team, T., 2017). The Bayes Factor varies as a function of the set prior. In our analyses we always set the Cauchy prior width at 0.707, and in Appendix B we show how the Bayes Factor varies with different values of this prior. These plots show that, although the strength of the evidence varies with the prior, in most cases the direction of the Bayes Factor is not affected by our choice of prior.

For unimodal test items there was no difference between items studied unimodally and items studied with a congruent sound, $t(50) = 1.67, p = .101, BF_{01} = 1.79$.¹ For unimodal test items there was also no difference between items studied unimodally and items studied with an

incongruent sound, $t(50) = 1.63, p = .109, BF_{01} = 1.90$. Plots showing the Bayes Factor as a function of the Cauchy prior width are provided in Appendix B.

Thelen et al. (2015) found that, compared to memory for items studied as pictures only, memory for items studied with congruent sounds was better and memory for items studied with incongruent sounds was worse. In contrast, the present experiment did not show such effect. If anything, our results indicated better memory for items studied as picture only than for items studied with a sound.

2. Experiment 2

In Experiment 2 memory for sounds was investigated. Participants studied sounds with or with a picture in a continuous recognition paradigm as was used in Experiment 1.

2.1. Method

2.1.1. Participants

Sixty-six students at the Erasmus University participated for course credit.

2.1.2. Stimuli and procedure

The same set of stimuli was used as in Experiment 1. Each sound was presented twice, and could be presented as sound-only, with a congruent picture, or with an incongruent picture. The procedure was the same as that used for Experiment 1, with the exception that participants responded whether the sound was 'old' or 'new'. They were informed

¹ The subscript indicated whether the data are more likely under the null hypothesis (BF_{01}) or the alternative hypothesis (BF_{10}).

that some of the sounds were presented together with a picture, but that the old-new decision should be based only on the sound.

2.2. Results

Data from 4 participants were excluded because the sounds did not play properly during their session, and data from 13 participants were excluded because their accuracy was below 60%. Only trials on which a participant responded were included (91.7% of all trials). The mean hit rates and false alarm rates are presented in Table 2. The average d 's for all conditions are shown in Fig. 2. A 3 (study condition) by 3 (test condition) ANOVA showed that memory was better for items studied with pictures than as sound-only, $F(2,96) = 9.01, p < .001$, partial $\eta^2 = 0.16$, and this effect seems to be restricted to the items that were tested with the same picture (i.e., the picture with which it had been presented on the first presentation), as indicated by the interaction effect, $F(4,192) = 29.27, p < .001$, partial $\eta^2 = 0.38$. Test format did not affect memory, $F(2,96) = 1.09, p = .341$, partial $\eta^2 = 0.02$.

To compare our results with those of Thelen et al., we analyzed the sound-only test condition separately and, compared to items studied as unimodal, found no advantage for items studied as bimodal-congruent, $t(48) = 0.49, p = .626, BF_{01} = 5.74$, nor a disadvantage for items studied as bimodal-incongruent, $t(48) = 0.51, p = .614, BF_{01} = 5.70$. Thus, we did not replicate Thelen et al.'s results.

Overall, these results suggest that task-irrelevant pictures may support memory for sounds, but only when the picture during study and test was the same. Interestingly this was found for both congruent and incongruent pictures. In other words, it was the overlap between study and test that mattered, not the congruency between the sound and the picture.

It is surprising that we did not replicate Thelen et al.'s (2015) findings of better memory for unimodal test items that had been studied as bimodal-congruent items than for unimodal test items studied as unimodal items. A possible explanation is that participants did not have enough time or resources to process the task-irrelevant modality. The presentation speed of items in the continuous recognition task was quite fast. The time from onset of one stimulus to the onset of the next stimulus varied between 1400 and 2000 ms, during which time participants had to process the stimulus, decide whether it was first presented or repeated, and make a response. Under such demanding conditions when there is competition for resources, top-down processes might work to select the task-relevant modality and ignore the task-irrelevant modality (Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010, but see Alais, Morrone, & Burr, 2006). The absence of a multisensory study advantage in our results might be due to the demanding task that caused participants to ignore the task-irrelevant modality. Although we used the same timing of presentations as Thelen et al. they presented the same items in a picture recognition and a sound recognition task to the same participants, which may have increased their participants' attention to the task-irrelevant modality despite the demanding conditions. The fast pacing is not critical to finding a multisensory benefit. To accommodate fMRI measurements, Murray et al. (2005) varied the intertrial interval between 6000 and 10,000 ms and also observed a multisensory benefit. In the next two experiments we increased the time between trials with 2000 ms in an attempt to make the task less demanding and give more room for attention to the task-irrelevant modality.

3. Experiment 3

3.1. Method

3.1.1. Participants

Sixty students at the Erasmus University participated for course credit.

3.1.2. Stimuli and procedure

The stimuli and procedure were the same as those in Experiment 1, except that the ISI was increased by 2000 ms to a variable duration between 2900 ms and 3500 ms.

3.2. Results

The accuracy for new and old items was calculated for each condition. Data from four participants were removed from the analyses because their accuracy was below 60%. Only trials on which a participant responded were included (92.4% of all trials). The mean hit rates and false alarm rates are presented in Table 2. The average d 's for all conditions are shown in Fig. 3. A 3 (study condition) by 3 (test condition) ANOVA showed no effects of study condition, $F(2,110) = 1.40, p = .251$, partial $\eta^2 = 0.03$, or test condition, $F(2,110) = 0.24, p = .789$, partial $\eta^2 = 0.00$. There was no interaction between study and test condition, $F(4,220) = 1.66, p = .162$, partial $\eta^2 = 0.03$. The absence of an interaction indicates that the overlap between study and test format did not affect memory.

A separate analysis of the items tested as picture-only showed, compared to items studied as unimodal, no advantage for items studied as bimodal-congruent, $t(55) = 0.09, p = .931, BF_{01} = 6.83$, nor a disadvantage for items studied as bimodal-incongruent, $t(55) = 0.93, p = .357, BF_{01} = 4.55$. Thus, increasing the inter-stimulus interval did not result in an effect of study modality on recognition memory performance for pictures.

4. Experiment 4

4.1. Method

4.1.1. Participants

Sixty-six students at the Erasmus University participated for course credit.

4.1.2. Stimuli and procedure

The stimuli and procedure were the same as those in Experiment 2, except that the ISI was increased by 2000 ms to a variable duration between 2900 and 3500 ms as in Experiment 3.

4.2. Results

Data from seven participants were excluded because their accuracy was below 60%. Only trials on which a participant responded were included (92.6% of all trials). The mean hit rates and false alarm rates are presented in Table 2. The average d 's for all conditions are shown in Fig. 4. A 3 (study condition) by 3 (test condition) ANOVA showed that memory was better for items studied with pictures than as sound-only, $F(2,116) = 31.42, p < .001$, partial $\eta^2 = 0.35$, and this effect seems to be restricted to the items that were tested with a picture and mainly due to better memory when the picture during study and test was the same, as indicated by the interaction effect, $F(4,232) = 26.27, p < .001$, partial $\eta^2 = 0.31$. In addition, there was an effect of test condition, memory was better for items tested with a picture than as sound-only, $F(2,116) = 6.22, p = .003$, partial $\eta^2 = 0.10$.

A separate analysis of the items tested as sound-only showed, compared to items studied as unimodal, an advantage for items studied as bimodal-congruent, $t(55) = 2.25, p = .028$, although the Bayesian analysis indicates that the evidence was very weak, $BF_{10} = 1.45$. We found no disadvantage for items studied as bimodal-incongruent, $t(58) = 1.55, p = .125, BF_{01} = 2.26$. Thus, we partially replicated Thelen et al. (2015) who found better memory for sounds studied with a congruent picture than items studied as sound only, although the Bayesian analysis indicated that the evidence was very weak. We did not replicate their finding that memory was better for items studied as sound-only than items studied with an incongruent picture.

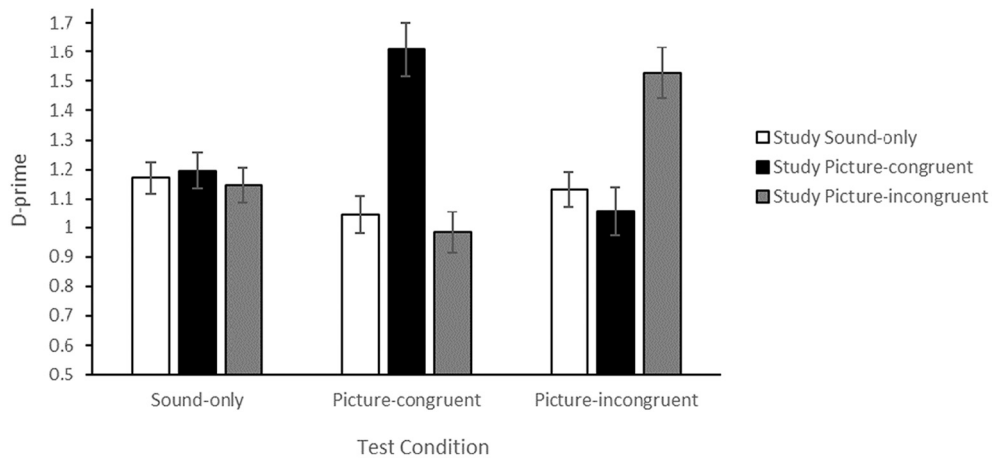


Fig. 2. Average D-primes in the sound recognition memory task of experiment 2 (error bars represent standard errors of the mean).

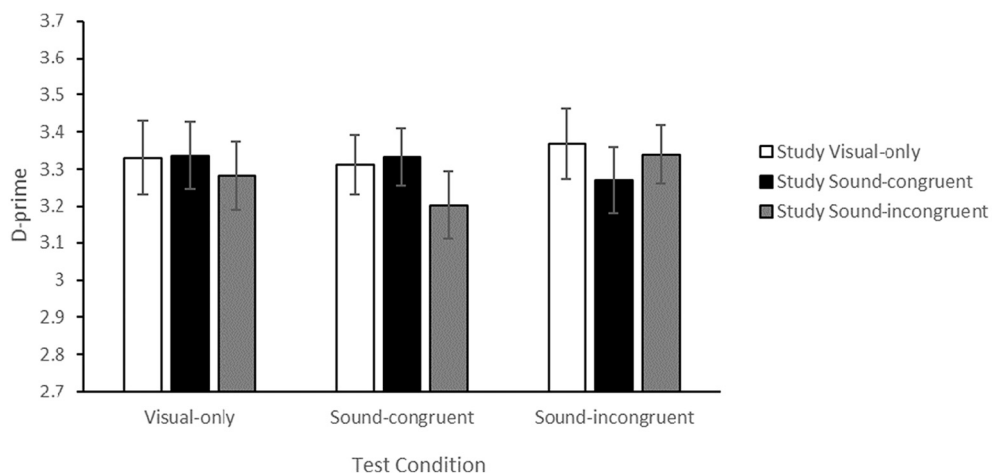


Fig. 3. Average D-primes in the picture recognition memory task of experiment 3 (error bars represent standard errors of the mean).

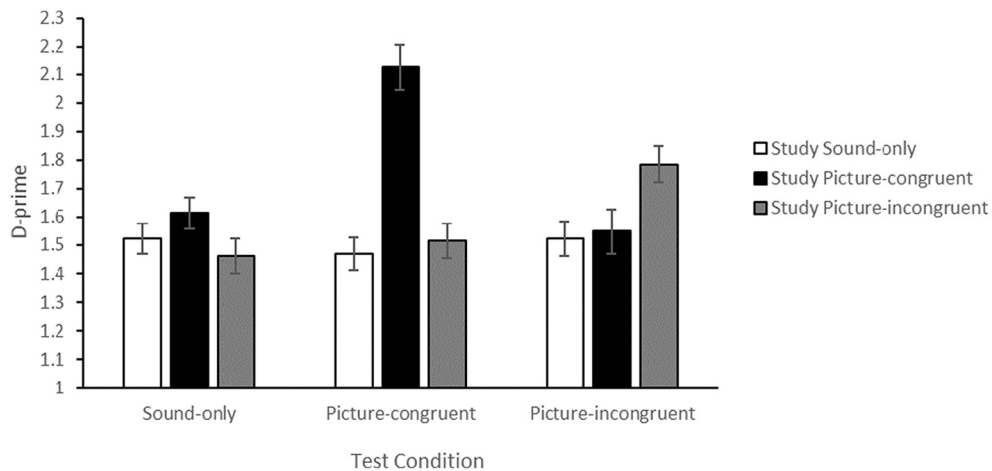


Fig. 4. Average D-primes in the sound recognition memory task of experiment 4 (error bars represent standard errors of the mean).

5. General discussion

In four experiments we tested memory for pictures (Experiments 1 and 3) and sounds (Experiments 2 and 4) representing common objects and animals. Items were presented as unimodal, bimodal-congruent, and bimodal-incongruent items in a continuous recognition task.

Format for new and repeated items could be the same or different in a fully crossed design. The results showed little evidence for a multisensory benefit and some evidence that memory for sounds benefited from study-test overlap. Based on the results of Thelen et al. (2015) and others (Lehmann & Murray, 2005; Moran et al., 2013), we had expected an overall benefit for bimodal-congruent study conditions over unimodal

study conditions. However, as shown in Figs. 1–4 no such omnipresent benefit was obtained; memory performance in the bimodal-congruent study condition was not consistently higher than in the unimodal study condition. A benefit for bimodal-congruent items was found, but only present when participants had to memorize sounds (see Figs. 2 and 4) and there largely restricted to bimodal-congruent test items that had also been studied as bimodal-congruent items. Importantly, and contrary to the idea that memory is only enhanced for bimodal-congruent items, a similar pattern was found for bimodal-incongruent items, indicating that these results are due study-test overlap. In the unimodal test condition in particular we had expected performance to be better when items had been studied in the bimodal-congruent condition than when items had been studied in the unimodal condition. In Figs. 1–4, performance in the unimodal test condition is shown in the three left-most bars. As can be seen, we did not find an advantage for the bimodal-congruent study condition over the unimodal study condition, except a very weak effect in Experiment 4. In addition, we also did not find that performance in the bimodal-incongruent study condition was below that of the unimodal study condition.

We were also interested in possible effects of study-test overlap in modality. Based on the encoding specificity and transfer-appropriate processing principles we expected better performance if the format of items (unimodal, bimodal-congruent, bimodal-incongruent) were the same during study and test, compared to when they were different (see also Meyerhoff & Huff, 2016). When memory was tested for pictures, the presence of a sound during study did not have an effect, nor was there an effect of study-test overlap (see Figs. 1 and 3). When memory was tested for sounds, however, there was an effect of the presence of a picture and this effect was primarily due to study-test overlap (see Figs. 2 and 4). We even found that memory performance for sounds was improved by a semantically incongruent picture at study if the same incongruent picture was also presented at test. Sound memory might be affected more by a picture than picture memory by a sound because memory for sounds is weaker than memory for pictures (Cohen et al., 2009; Heikkilä et al., 2017). Moreover, visual information is dominant in general (Colavita, 1974), also in processing meaningful pictures and sounds (Yuval-Greenberg & Deouell, 2009). This means that task-irrelevant visual information is harder to ignore than task-irrelevant auditory information. Therefore, sound memory may have benefitted from the overlap in visual information between study and test because sound memory is weak and visual information is hard to ignore. Picture memory, on the other hand, may not have benefitted from overlap in sound information because picture memory is strong already and sound information is easy to ignore. Thus, overlap in sound would be more effective than overlap in picture.

The current study used d' as a measure of recognition memory accuracy (Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988). Although this is the standard measure of performance in recognition memory tasks, and much preferred over just considering hit rates, d' accurately reflects recognition memory accuracy only if the assumptions underlying its calculation are met. One such assumption is that the variance of the target and foil distributions of familiarity are equal. Studies indicate, however, that this equal-variance assumption is consistently violated (e.g., Ratcliff, Gronlund, & Sheu, 1992); the variance of the target distribution is typically larger than the variance of the foil distribution. The violation of the equal-variance assumption is particularly problematic when comparing conditions in which participants use different response biases (e.g., Grider & Malmberg, 2008). As we argued in the Introduction, such differences may have been present in studies where there was a correlation between presentation format (unimodal vs. bimodal) and study status (studied vs. nonstudied). The present study, however, was designed so that there was no correlation between presentation format and study status, thereby eliminating the motivation for participants to be differently biased in the unimodal and bimodal conditions. To the extent that in the present study d' accurately reflected differences in performance between conditions, we would

expect to find similar findings for alternative measures of memory accuracy, such as d'_e (Grider & Malmberg, 2008),² or with procedures developed to eliminate response biases, such as two-alternative forced-choice (Grider & Malmberg, 2008; Zeelenberg, Wagenmakers, & Raaijmakers, 2002; Zeelenberg, Wagenmakers, & Rotteveel, 2006). To our knowledge, no study on multisensory memory has used such alternative approaches.

The lack of a general multisensory memory benefit in our study suggests that the effect is not robust and may be sensitive to task-specific factors. One likely factor is that the presence of a multisensory benefit depends on the amount of attention that participants pay to the task-irrelevant modality. In Thelen et al. (2015), all participants were tested in both a picture recognition memory and a sound recognition memory task (see also Heikkilä et al., 2015; Heikkilä et al., 2017). Even though Thelen et al. separated these tasks by a week, participants may still have paid attention to the task-irrelevant modality in the second task, especially because the same target items were presented in both tasks and the task-irrelevant modality in the second task had been the task-relevant modality in the first task. For example, if the second task was sound memory, bimodal items would be accompanied by pictures that had been the to-be-remembered items in the first task and may for that reason have captured attention.

Such a task-specific carry-over effect on attention cannot explain the multisensory benefit obtained by three other studies in which only one modality was task-relevant (Lehmann & Murray, 2005; Moran et al., 2013; Murray et al., 2005). In the studies by Lehmann and Murray (Experiment 2) and Murray et al. participants had to remember pictures while sounds were task-irrelevant, in the study by Moran et al. participants had to remember sounds while pictures were task-irrelevant. That is, in contrast of the Thelen et al. (2015) study, only one modality was ever task relevant. In these studies, however, items were presented in a continuous recognition task as either unimodal or bimodal items on their first presentation, but always as unimodal items on their second presentation. As discussed in the Introduction, this may have introduced a response bias based on the modality of the items. Moreover, because across all trials the correct response to a bimodal item was always 'new' and the correct response to a unimodal item was more likely to be 'old' than 'new', participants have been motivated to pay close attention to the task-irrelevant modality because that would help them make their responses. In addition, in this design unimodal items are more frequent (75%) than bimodal items (25%). Unusual items are commonly better remembered than usual items when the two types are mixed (McDaniel & Bugg, 2008). Because the bimodal items were the unusual items it is possible that the memory benefit for bimodal items was the result of their unusualness rather than the bimodal format itself. To circumvent these potential problems, we used the continuous recognition task with the design of Thelen et al. (2015) in which all formats were equally likely. Our results suggest that using such a design eliminates the multisensory memory benefit.

We are not the first study to obtain little evidence for a multisensory memory benefit. Studies that presented items in separate study and test blocks also did not obtain a multisensory advantage (Cohen et al., 2009; Nyberg et al., 2000). Although this is a different procedure than the continuous recognition task there is no obvious reason why the effect of multimodal presentation format should be restricted to the continuous recognition task. In fact, Duarte, Ghetti, and Geng (2021) have suggested that the multisensory advantage might be limited to recollection processes in memory. Such recollection processes might play a larger role when participants have more time to respond as might be the case in blocked study-test designs compared to the fast-paced continuous recognition paradigm. In both procedures participants receive explicit memory instructions and the test taps into long-term memory. These and

² Note that the calculation of d'_e requires confidence ratings which we did not collect in our study.

our results suggest that the multisensory memory benefit is not a general finding but rather one that is restricted to specific task details.

A multisensory benefit in memory would be consistent with findings of multisensory benefits in attention (Evans, 2020; Quak et al., 2015; Talsma et al., 2010). Attention is increased by items in a task-irrelevant modality that are spatially or temporally aligned with the task-relevant stimulus. For example, Van der Burg, Olivers, Bronkhorst, and Theeuwes (2008) found that participants were faster to detect the orientation of a horizontal or vertical line segment among oblique distractor line segments if a change in color of the target segment was accompanied by a tone than if the color change occurred without a tone. The stimuli used in studies that investigated the effect of multimodal stimuli on attention were often simple stimuli such as visual gratings or single tones. The effects of unimodal vs multimodal presentations on attentional processes with more complex stimuli such as a picture of a dog or the sound of a dog bark might be different because they require deeper processing to establish congruency. Thus, processing the second, task-irrelevant modality may actually hurt performance because it requires processing resources and thus takes attention away from the task-relevant modality. Moreover, evidence suggests that conditions that facilitate initial processing of a stimulus may not always result in better memory. For example, Canits et al. (2018) found no benefit of congruent motor actions on later memory for pictures even though congruency did affect immediate motor responses during the study task. Thus, although it may seem reasonable to expect memory benefits for congruent conditions that show an effect during study, the relation between congruency during study and later memory performance may not be so straightforward.

The effect of study-test overlap in Experiments 2 and 4 shows that participants did not completely ignore the task-irrelevant modality. It appears that some information from the task-irrelevant modality was encoded in memory, otherwise the overlap between study and test would not have affected memory performance. However, the irrelevant

information at study was not helpful unless it was also presented at test. A similar conclusion can be drawn from Nyberg et al. (2000) who measured activation in auditory brain regions during a memory test for pictures. They showed that these auditory regions were activated more by unimodal test pictures that had been studied with a sound than by pictures that had been studied without a sound. Responses were not more accurate, however, for items studied bimodally than for items studied unimodally. Thus, even though these results suggest that irrelevant sounds were encoded at study and activated during test they did not support memory performance.

To conclude, our results show little benefit of multisensory study on memory. Previous findings of such a benefit may have been due to context-specific carry-over effects of attention allocation or to unequal distributions of items over conditions and responses. Our results showed that only when the exact multisensory item is repeated at test a strong memory benefit due to study-test overlap is present for both congruent and incongruent items, but this does not indicate a general multisensory integration benefit on memory performance.

Author note

We thank Corinna Börner, Jona Frasch, Margriet Meijers, Sana Oulad El Arbi, Benjamin Quist, Sylvia Reverda, Tapio Somersalbi, Justine Staal, Nian Tang, and the staff at the EBL (Gerrit Jan de Bie, Christiaan Tieman, and Marcel Boon) for their help on this project. We have no competing interests to declare.

CRedit authorship contribution statement

Diane Pecher: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Supervision. **René Zeelenberg:** Conceptualization, Methodology, Formal analysis, Writing – review & editing.

Appendix A

Experimental Items

trumpet, leopard, bee, saxophone, telephone, teakettle, brush, chicken, rhino, sheep, tiger, helicopter, chainsaw, car, canary, goat, donkey, rooster, jetplane, scissors, kitchen stove, grizzly bear, fire extinguisher, guitar, wasp, accordion, songbird, drill, typewriter, whale, swan, tennis, french horn, dolphin, ambulance, hammer, piano, lawnmower, violin, microwave, cannon, seagull, sewing machine, camel, computer mouse, sword, dog, duck, photocopier, truck, cow, pencil, cash register, hairdryer, elephant, knife, motorcycle, soda can, treadmill, eagle, dynamite, axe, police car, motorboat, tambourine, mobile phone, broom, door, book, harmonica, rifle, sleigh bells, torch, saw, squirrel, monkey, toaster, fly, locking, watercooler, saber, heart monitor, oceanliner, cricket, frying pan, gong, water tap, champagne, clock, slot machine, bat, train, horse, sink, mouse, pingpong bat, dice, owl, gun, raccoon, vacuum cleaner, gavel, record, washer, alarm, tractor, stethoscope, stapler, hyena, hockey stick, choir, electric saw, firetruck, aerosol can, biplane, plank, campfire, chisel, plunger, lion, parrot, crocodile, pencil sharpener, race car, rake, drum, lamb, tank, zebra, coffeemaker, light saber, ball, clown, submarine, whistle, anchor, wolf, xylophone, golf stick, manometer, kettle, school bell, snake, cheering audience, broken bottle, harp, bow and arrow, rocket, clapping hands, railroad crossing, chain, triangle, boiling water, banjo, camera, gorilla, shaver, raven, penguin, sealion, blender, keyboard, peacock, key, light switch, robot, chalkboard, guillotine, cat, radio, saltshaker, toilet, bell, flag, hippopotamus, toothbrush, dinosaur, fan, drinking straw, maracas, pill bottle, boot spurs, tornado, bowling pin, pig, church, grenade, frog, skateboard, printer, shoe, shopping cart, tape dispenser, nail clippers, rain, guinea pig, top, football, fountain, didgeridoo, castanets, balloon, mosquito, swimming pool, broken cup, flute, apple bite, lighter, baseball mitt, potato chips, witch, boxing glove, nose, billiards, elevator, bath, torn paper, zipper, marker, coin, baby, ghost, lips, latex gloves, corkscrew, rubber duck, metronome, kettle, fireworks, cell phone, alarm clock.

Appendix B

Plots Showing Bayes Factors as a Function of Cauchy Prior Width for Analyses on Unimodal Test Items

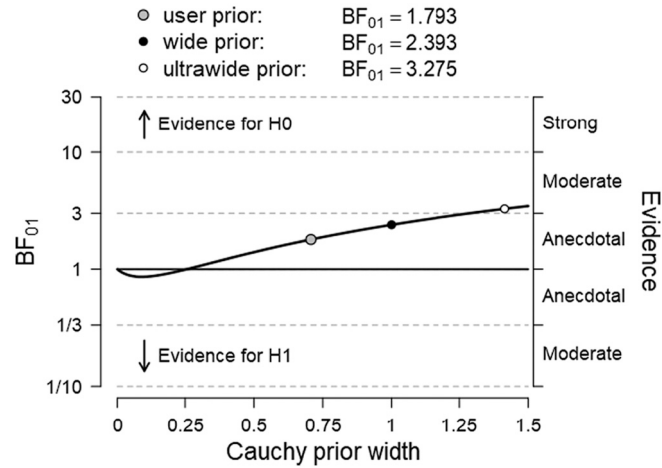


Fig. B1. Bayes factors as a function of Cauchy prior width for experiment 1 (unimodal study condition vs. bimodal-congruent study condition).

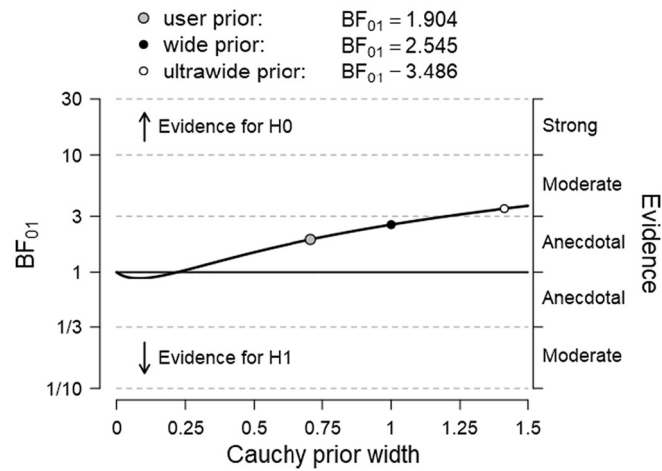


Fig. B2. Bayes factors as a function of Cauchy prior width for experiment 1 (unimodal study condition vs. bimodal-incongruent study condition).

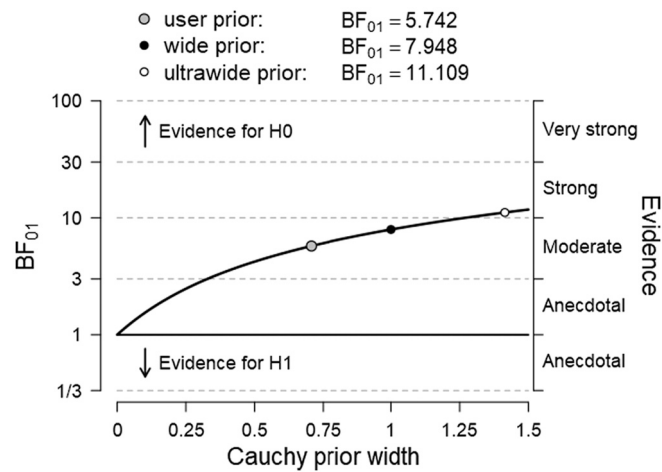


Fig. B3. Bayes factors as a function of Cauchy prior width for experiment 2 (unimodal study condition vs. bimodal-congruent study condition).

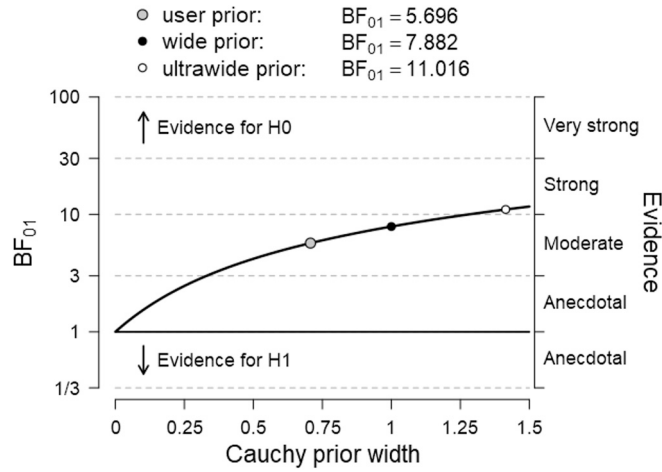


Fig. B4. Bayes factors as a function of Cauchy prior width for experiment 2 (unimodal study condition vs. bimodal-incongruent study condition).

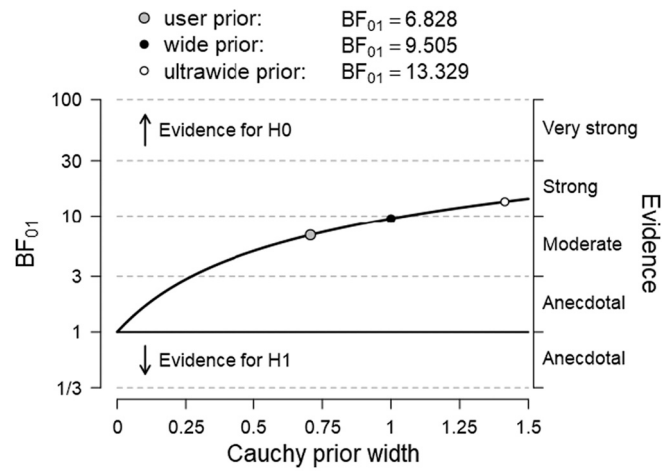


Fig. B5. Bayes factors as a function of Cauchy prior width for experiment 3 (unimodal study condition vs. bimodal-congruent study condition).

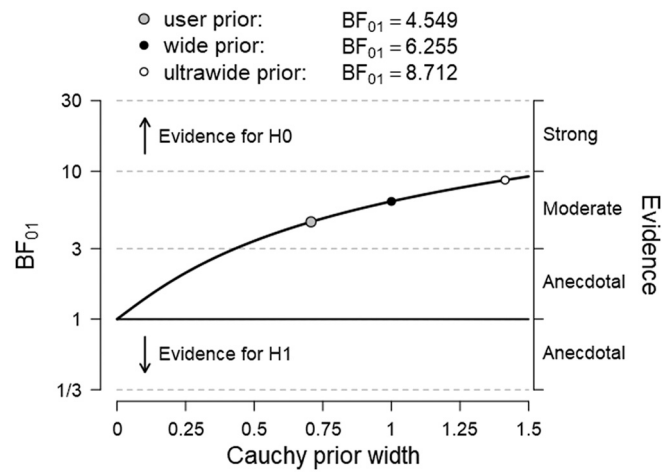


Fig. B6. Bayes factors as a function of Cauchy prior width for experiment 3 (unimodal study condition vs. bimodal-incongruent study condition).

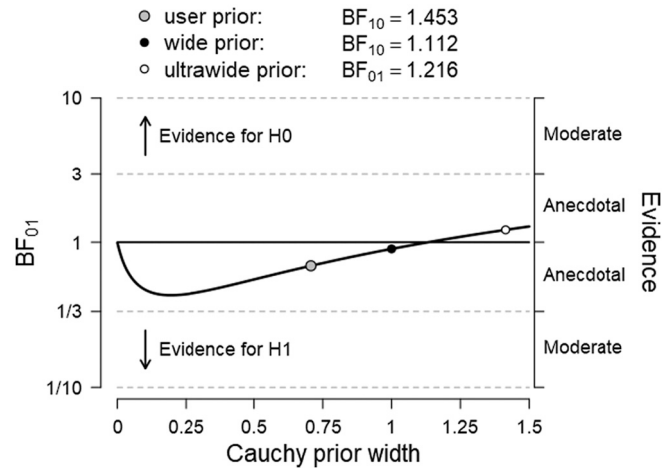


Fig. B7. Bayes factors as a function of Cauchy prior width for experiment 4 (unimodal study condition vs. bimodal-congruent study condition).

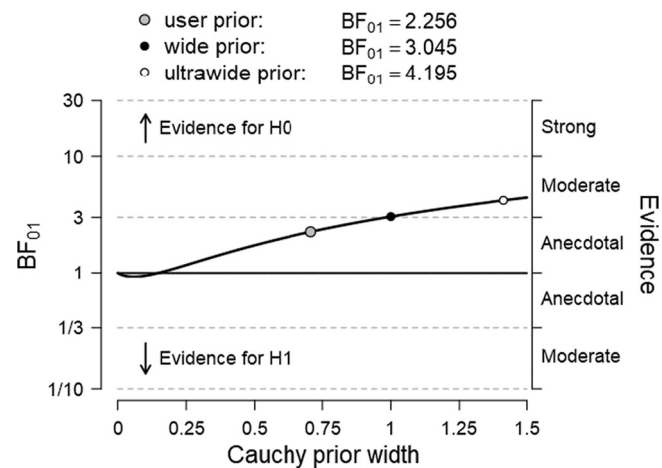


Fig. B8. Bayes factors as a function of Cauchy prior width for experiment 4 (unimodal study condition vs. bimodal-incongruent study condition).

References

- Alais, D., Morrone, C., & Burr, D. (2006). Separate attentional resources for vision and audition. *Proceedings of the Royal Society B-Biological Sciences*, 273(1592), 1339–1345. <https://doi.org/10.1098/rspb.2005.3420>
- Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrel, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, 13(4), 471–481. [https://doi.org/10.1016/S0022-5371\(74\)80024-1](https://doi.org/10.1016/S0022-5371(74)80024-1)
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15(11), 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(4), 657–668.
- Cabeza, R. (1994). A dissociation between two implicit conceptual tests supports the distinction between types of conceptual processing. *Psychonomic Bulletin & Review*, 1(4), 505–508.
- Canits, I., Pecher, D., & Zeelenberg, R. (2018). Effects of grasp compatibility on long-term memory for objects. *Acta Psychologica*, 182, 65–74. <https://doi.org/10.1016/j.actpsy.2017.11.009>
- Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America*, 106(14), 6008–6010. <https://doi.org/10.1073/pnas.0811884106>
- Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, 16(2), 409–412.
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33(1–2), 25–62. [https://doi.org/10.1016/0010-0277\(89\)90005-X](https://doi.org/10.1016/0010-0277(89)90005-X)
- Duarte, S. E., Ghetti, S., & Geng, J. (2021). Object memory is multisensory: Task-irrelevant sounds improve recollection-based recognition memory. <https://doi.org/10.31234/osf.io/pk4cf>. September 21.
- Evans, K. K. (2020). The role of selective attention in cross-modal interactions between auditory and visual features. *Cognition*, 196, 104119. <https://doi.org/10.1016/j.cognition.2019.104119>
- Grider, R. C., & Malmberg, K. J. (2008). Discriminating between changes in bias and changes in accuracy for recognition memory of emotional stimuli. *Memory & Cognition*, 36(5), 933–946. <https://doi.org/10.3758/MC.36.5.933>
- Heikkilä, J., Alho, K., Hyvönen, H., & Tiippana, K. (2015). Audiovisual semantic congruency during encoding enhances memory performance. *Experimental Psychology*, 62(2), 123–130. <https://doi.org/10.1027/1618-3169/a000279>
- Heikkilä, J., Alho, K., & Tiippana, K. (2017). Semantically congruent visual stimuli can improve auditory memory. *Multisensory Research*, 30(7–8), 639–651. <https://doi.org/10.1163/22134808-00002584>
- Heikkilä, J., & Tiippana, K. (2016). School-aged children can benefit from audiovisual semantic congruency during memory encoding. *Experimental Brain Research*, 234(5), 1199–1207. <https://doi.org/10.1007/s00221-015-4341-6>
- Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, 32(4), 421–445. <https://doi.org/10.1006/jmla.1993.1023>
- JASP Team, T. (2017). *JASP (version 0.8.1.2) [computer software]*.
- von Kriegstein, K., & Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biology*, 4(10), 1809–1820. <https://doi.org/10.1371/journal.pbio.0040326>
- Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Cognitive Brain Research*, 24(2), 326–334. <https://doi.org/10.1016/j.cogbrainres.2005.02.005>
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Lawrence Erlbaum Associates Publishers.
- 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*, 41(5), 699–708. <https://doi.org/10.1111/ejn.12804>

- Mayer, R. E. (Ed.). (2008). *The Cambridge handbook of multimedia learning*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.002>.
- McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin and Review*, 15(2), 237–255. <https://doi.org/10.3758/PBR.15.2.237>
- Meyerhoff, H. S., & Huff, M. (2016). Semantic congruency but not temporal synchrony enhances long-term memory performance for audio-visual scenes. *Memory & Cognition*, 44(3), 390–402. <https://doi.org/10.3758/s13421-015-0575-6>
- Moran, Z. D., Bachman, P., Pham, P., Cho, S. H., Cannon, T. D., & Shams, L. (2013). Multisensory encoding improves auditory recognition. *Multisensory Research*, 26(6), 581–592. <https://doi.org/10.1163/22134808-00002436>
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 519–533.
- Murray, M. M., John, J. F., & Wylie, G. R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage*, 27(2), 473–478. <https://doi.org/10.1016/j.neuroimage.2005.04.016>
- Murray, M. M., Michel, C. M., de Peralta, R., 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. <https://doi.org/10.1016/j.neuroimage.2003.09.035>
- Nyberg, L., Habib, R., McIntosh, A., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20), 11120–11124. <https://doi.org/10.1073/pnas.97.20.11120>
- Parks, C. M. (2013). Transfer-appropriate processing in recognition memory: Perceptual and conceptual effects on recognition memory depend on task demands. *Journal of Experimental Psychology-Learning Memory and Cognition*, 39(4), 1280–1286. <https://doi.org/10.1037/a0030911>
- Pecher, D., Zanolie, K., & Zeelenberg, R. (2007). Verifying visual properties in sentence verification facilitates picture recognition memory. *Experimental Psychology*, 54(3), 173–179. <https://doi.org/10.1027/1618-3169.54.3.173>
- Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, 11(1), 164–167. <https://doi.org/10.3758/BF03206477>
- Quak, M., London, R. E., & Talsma, D. (2015). A multisensory perspective of working memory. *Frontiers in Human Neuroscience*, 9, 197. <https://doi.org/10.3389/fnhum.2015.00197>
- Ratcliff, R., Gronlund, S. D., & Sheu, C. F. (1992). Testing global memory models using ROC curves. *Psychological Review*, 99(3), 518–535. <https://doi.org/10.1037/0033-295X.99.3.518>
- Roediger, H., & Blaxton, T. A. (1987). Effects of varying modality, surface-features, and retention interval on priming in word-fragment completion. *Memory & Cognition*, 15(5), 379–388. <https://doi.org/10.3758/BF03197728>
- Roediger, H. L., & Adelson, B. (1980). Semantic specificity in cued recall. *Memory & Cognition*, 8(1), 65–74. <https://doi.org/10.3758/bf03197553>
- Roediger, H. L., Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–39). Erlbaum.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review*, 16(2), 225–237. <https://doi.org/10.3758/PBR.16.2.225>
- Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, 12(11), 411–417. <https://doi.org/10.1016/j.tics.2008.07.006>
- Shepard, R. N., & Teghtsoonian, M. (1961). Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology*, 62(3), 302–309. <https://doi.org/10.1037/h0048606>
- Snodgrass, J., & Corwin, J. (1988). Pragmatics of measuring recognition memory - applications to dementia and amnesia. *Journal of Experimental Psychology-General*, 117(1), 34–50. <https://doi.org/10.1037/0096-3445.117.1.34>
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9), 400–410. <https://doi.org/10.1016/j.tics.2010.06.008>
- Thelen, A., & Murray, M. M. (2013). The efficacy of single-trial multisensory memories. *Multisensory Research*, 26(5), 483–502. <https://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. <https://doi.org/10.1016/j.cognition.2015.02.003>
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 359–380.
- Van Dantzig, S., Cowell, R. A., Zeelenberg, R., & Pecher, D. (2011). A sharp image or a sharp knife: Norms for the modality-exclusivity of 774 concept-property items. *Behavior Research Methods*, 43(1), 145–154. <https://doi.org/10.3758/s13428-010-0038-8>
- Van der Burg, E., Olivers, C., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology-Human Perception and Performance*, 34(5), 1053–1065. <https://doi.org/10.1037/0096-1523.34.5.1053>
- Woods, A., & Newell, F. (2004). Visual, haptic and cross-modal recognition of objects and scenes. *Journal of Physiology-Paris*, 98(1–3), 147–159. <https://doi.org/10.1016/j.jphysparis.2004.03.006>
- Yuval-Greenberg, S., & Deouell, L. Y. (2009). The dog's meow: Asymmetrical interaction in cross-modal object recognition. *Experimental Brain Research*, 193(4), 603–614. <https://doi.org/10.1007/s00221-008-1664-6>
- Zeelenberg, R. (2005). Encoding specificity manipulations do affect retrieval from memory. *Acta Psychologica*, 119(1), 107–121. <https://doi.org/10.1016/j.actpsy.2004.12.001>
- Zeelenberg, R., Pecher, D., Shiffrin, R. M., & Raaijmakers, J. G. W. (2003). Semantic context effects and priming in word association. *Psychonomic Bulletin & Review*, 10(3), 653–660. <https://doi.org/10.3758/BF03196528>
- Zeelenberg, R., Wagenmakers, E. J., & Rotteveel, M. (2006). The impact of emotion on perception. *Psychological Science*, 17(4), 287–291. <https://doi.org/10.1111/j.1467-9280.2006.01700.x>
- Zeelenberg, R., Wagenmakers, E. M., & Raaijmakers, J. G. W. (2002). Priming in implicit memory tasks: Prior study causes enhanced discriminability, not only bias. *Journal of Experimental Psychology: General*, 131(1), 38–47.