

Visuo-haptic integration in object identification using novel objects

Geneviève Desmarais¹ · Melissa Meade² · Taylor Wells¹ · Mélanie Nadeau¹

Published online: 25 July 2017
© The Psychonomic Society, Inc. 2017

Abstract Although some studies have shown that haptic and visual identification seem to rely on similar processes, few studies have directly compared the two. We investigated haptic and visual object identification by asking participants to learn to recognize (Experiments 1, and 3), or to match (Experiment 2) novel objects that varied only in shape. Participants explored objects haptically, visually, or bimodally, and were then asked to identify objects haptically and/or visually. We demonstrated that patterns of identification errors were similar across identification modality, independently of learning and testing condition, suggesting that the haptic and visual representations in memory were similar. We also demonstrated that identification performance depended on both learning and testing conditions: visual identification surpassed haptic identification only when participants explored the objects visually or bimodally. When participants explored the objects haptically, haptic and visual identification were equivalent. Interestingly, when participants were simultaneously presented with two objects (one was presented haptically, and one was presented visually), object similarity only influenced performance when participants were asked to indicate whether the two objects were the same, or when participants had learned about the objects visually—without any haptic input. The results suggest that haptic and visual object representations rely on similar processes, that

they may be shared, and that visual processing may not always lead to the best performance.

Keywords Multisensory processing · Object recognition · Haptics

Imagine searching through a bag for a specific item—your cell phone. This is an item you have learned to recognize by holding it, looking at it, and using it. Chances are you can recognize it by sight (visually) or by touch (haptically). How did the visual and haptic exposure shape your memory representation for this object? How did it impact your ability to recognize it? How do other items you see in the bag facilitate or interfere with your ability to identify and find your cell phone? Past research suggests that visual and haptic object identification are both effective, that they may rely on similar representations in memory, and that they follow similar principles—both seem to rely on shape perception. However, few studies have directly compared haptic and visual object identification, and how the two modalities interact when we recognize objects. In a first step, we investigated whether haptic and visual object representations in memory were comparable by evaluating whether haptic and visual object identification produced similar patterns of errors using a set of psychophysically scaled objects. We then examined the interaction between the two representations by presenting incongruent haptic and visual information and evaluated whether the similarity between the two inputs affected participants' responses.

✉ Geneviève Desmarais
gdesmarais@mta.ca

¹ Department of Psychology, Mount Allison University, 49A York Street, Sackville, NB E4L 1C7, Canada

² Department of Psychology, University of Waterloo, Waterloo, ON N2L 3G1, Canada

Haptic and visual object identification

Early research suggested that haptic processing was less accurate than visual processing (Bryant & Raz, 1975). However,

more recent studies have demonstrated that touch can be as accurate as vision for object identification. Klatzky, Lederman, and Metzger (1985) asked participants to name common objects that could be held within the hands (e.g., comb, whistle, mitten, etc.), and reported accuracy rates of 96%. This figure is comparable to those reported for visual identification by Lloyd-Jones and Humphreys (1997a): these authors observed accuracy rates between 86% (for visually similar objects) and 96% (for visually distinct objects). Additionally, Newell, Woods, Mernagh, and Bühlhoff (2005) have also observed similar error rates across visual and haptic scene recognition. The accuracy rates between visual and haptic identification can therefore be considered comparable.

Visual and haptic identification also seem to rely at least partially on shape information. Klatzky et al. (1985) asked participants to report the properties they had used to haptically identify objects, and observed that ‘global shape’ was used most often (46% of trials), followed by global texture (36% of trials) and the presence of a distinct component (35% of trials). Klatzky et al.’s observation was unfortunately subjective, as participants evaluated their own classification process, and were limited to haptic perception. More recently, Gaissert & Wallraven (2012) asked participants to classify seashells into groups using vision or touch. They observed that participants used comparable principles for visual and haptic classification. In both modalities, shape similarity was primarily used to form categories.

Consistent with evidence suggesting that individuals rely on comparable dimensions to classify the same objects visually or haptically, patterns of errors across dimensions are analogous. In a series of papers, Loomis (1981; 1982) has demonstrated that patterns of errors (specifically which two Braille letters or Roman characters are confused) are similar across the visual and haptic modalities. These similarities in error patterns led other researchers to suggest that visual processing and haptic processing shared overlapping memory representations.

Representation in memory for haptic and visual information

A number of different paradigms demonstrate that haptics and vision seem to share memory representations. Probing working memory, Norman, Norman, Clayton, Lianekhammy and Sielke (2004) asked participants to immediately match naturally shaped objects across the visual and haptic modalities. The authors observed a high accuracy rate across modalities (72%), as well as a slight advantage for matching items within the visual modality compared to cross-modal matching. Similarly, in an assessment of long-term episodic memory, Easton Greene, and Srinivas (1997) found that replication (drawing) and recognition of previously presented (old) and

new simple line patterns were superior for previously studied patterns, independently of whether the patterns were studied haptically or visually. Together, these findings suggest that haptic and visual representations in memory are shared.

However, the patterns of findings regarding the interaction of haptic and visual memory representations is not entirely consistent. First, Easton Greene, and Srinivas (1997) found that on the recognition test, participants consistently performed better when presented with line patterns they had studied visually, showing a clear advantage for visual relative to haptic encoding. But when real objects were used instead of three-line patterns, the accuracy of old/new judgments depended on both study and test conditions: visual recognition was better for visually studied objects, and haptic recognition was better for haptically presented objects. These latter observations are similar to those of Newell et al. (2005) where scene matching was more accurate when done within modality. As such, it seems apparent that there is some overlap in the way that haptic and visual information is represented in memory, though it is unclear to what extent these representations are shared, and whether there exists a bias for visual information.

Multimodal information

When information is presented simultaneously in more than one modality, there seems to be an advantage for visual information over haptic information. This has been observed for size estimation (Rock & Victor, 1964, but *see* Hershberger & Misceo, 1996 for contradictory findings), for reporting the modality of stimuli (Colavita, 1974; Hetch & Reiner, 2009), as well as with a number of multisensory illusions like the rubber-hand illusions (Botvinick & Cohen, 1998) and a visuo-haptic variant of the Müller-Lyer illusion (Mancini, Bricolo, & Vallar, 2010). Though a few studies have reported an advantage for haptic processing, this advantage tended to occur under conditions of degraded vision (Ernst & Banks, 2002; Ernst & Bühlhoff, 2004; Heller, 1983). Pavani, Spence, and Driver (2000) and Spence, Pavani, and Driver (2004) provided strong evidence for the influence of vision over touch, even when visual information is irrelevant. They asked participants to indicate the location of a vibro-tactile stimulus on a hand-held block, and concurrently presented an irrelevant distracter light at either the same location or a different location. Distractors at different locations interfered with vibro-tactile localization, even when the distractors were presented further away from the targets (Spence et al., 2004), and despite the fact that the visual stimuli were irrelevant to the task. Though the authors evaluated the impact of visual information on tactile perception, they did not investigate whether tactile information could impact visual perception. Nonetheless, even irrelevant visual information that is spatially separated

can bias the haptic perception of location. Altogether, though there is definitely evidence for visual dominance over haptic processing, few studies have looked at whether haptic information can bias visual identification, and none have done it in the context of object identification.

Real objects versus novel objects

Many of the studies that examined haptic identification or recognition used real objects (Gaissert & Wallraven, 2012; Klatzky et al., 1985). One issue that arises with the use of real objects is that it is not possible to know which object dimensions are used to make comparisons or identifications. Even though Gaissert & Wallraven (2012) and Klatzky et al. (1985) demonstrated that haptic and visual identification rely partially on form information, other kinds of information (e.g., color and lighting) are only available to vision, while others (e.g., surface features like stickiness) may only be available to touch, making the extent of the overlap between visual and haptic representations unclear. In order to get around these kinds of difficulties, Desmarais, Dixon, and Roy (2007) used novel objects that were psychophysically scaled (*see* Desmarais & Dixon, 2005) such that a change in one dimension (e.g., curvature) was equal to a change in another dimension (tapering or thickness). Importantly, these changes affected the form of the object—the perceptual element reported by Gaissert & Wallraven (2012) and Klatzky et al. (1985) as most important for visual and haptic object categorization. Desmarais et al., (2007) have also confirmed that, with this set of stimuli, similar objects are confused more often than distinct objects, replicating the pattern observed with real objects (*see* Dickerson & Humphreys, 1999; Humphreys, Riddoch and Quinan, 1988; Lloyd-Jones & Humphreys, 1997a, b).

Current study

In a first step (the learning phase of Experiment 1), we investigated whether visual and haptic identification are comparable, and whether performance depends on matching learning and testing modality. In a second step (the experimental phase of Experiment 1 and later in Experiments 2 and 3), we presented visual information and haptic information concurrently to investigate whether irrelevant information would interfere with identification, and whether this interference would be mediated by object similarity.

Participants first completed a learning phase. We presented three separate groups of participants with novel objects along with a nonword name. Participants were asked to explore the objects either haptically, visually, or bimodally, and were asked to remember the name of each

object. We then asked all participants to recognize the objects haptically and visually. First, we hypothesized that object similarity would drive object identification: for both haptic and visual identification, similar objects would be confused more often than distinct objects. Secondly, because performance has been better when testing conditions match learning conditions (Easton, Greene, & Srinivas, 1997; Hershberger & Misceo, 1996; Newell et al., 2005), we hypothesized that participants would perform better when their learning and testing conditions matched. Specifically, participants would require fewer blocks to reach flawless performance, and would produce fewer errors if learning and testing were both haptic or both visual. Finally, for participants in the bimodal learning condition, we predicted that performance would mirror that of participants in the visual learning condition: they would require fewer blocks to criterion and will produce fewer errors when tested visually than when tested haptically.

The same participants then completed an experimental phase where we presented them with two objects simultaneously (one was haptically presented, and one was visually presented), and asked them to identify one of the two objects. Crucially, in half of the trials, the two objects were copies of the same objects (congruent trials), while in the other half of the trials, the two objects were copies of two different objects (incongruent trials). Again, we expected participants to be faster and more accurate when learning and testing conditions matched. Second, we expected incongruent distractors to create interference and slow reaction times, but we hypothesized that this interference would be mediated by testing modality. Given the general predominance of vision over touch when visual and haptic information are simultaneous (Pavani et al., 2000; Spence et al., 2004), we expected a greater amount of interference during haptic identification than during visual identification. Finally, we expect interference to be modulated by similarity: we expect more interference from similar distractors than from distinct distractors.

Experiment 1

Method

Participants

Seventy undergraduate students (53 females, mean age = 19.44 years) were recruited from introduction to psychology courses at Mount Allison University. All participants confirmed being right-handed, and having normal or corrected-to-normal vision and hearing. Each participant was compensated with 0.5 course credit for each half hour of participation.

Materials

We used a set of eight novel objects (81 mm long, made of hard, light-gray PVC) that were psychophysically scaled by Desmarais and Dixon (2005). These objects vary in thickness, curvature, and tapering. These visual dimensions form a three-dimensional space where the position of each object can be described by a set of Cartesian coordinates, where each axis refers to a stimulus dimension (curvature, tapering, and thickness) and the coordinate ‘1’ refers to the presence of that dimension. In this space the object in position 0,0,0 is called the “origin object,” and the values indicate that it is not curved, is not tapered and has (for ease of labeling) minimal thickness. All other objects involve perturbations of this origin object. In this space, some objects are close together and some are further apart. The distance between two objects can be construed as city-block (CB) distance, where the distance between two objects differing on a given number of dimensions is defined as the sum of the distances obtained separately on each of the dimensions (Shepard, 1987 has demonstrated that shape perception is based on city-block metric rather than Euclidian metric). Note that the number of distinguishing attributes equals the number of CBs that separate two objects. Objects 0,0,0 and 0,0,1 are one CB apart, whereas objects 0,0,0 and 1,0,1 are two CBs apart, and objects 0,0,0 and 1,1,1 are three CBs apart. Desmarais and Dixon (2005) psychophysically scaled these objects to ensure that changes on each of the three dimensions were equally salient: changes in one dimension were perceived as equivalent to changes in another dimension. By using this object set, one knows, and can therefore control the visual similarity relations among the exemplars that make up this collection of objects.

A 45-cm × 40-cm × 44.5-cm black PVC apparatus was constructed (see Fig. 1) to allow participants haptic and/or visual perception of the objects. Within the apparatus are two identical sets of the eight novel objects on horizontal spokes attached to a central spinning cylinder. The objects of each set are positioned adjacent to one another with one set surrounding the cylinder near the bottom and the other set surrounding the cylinder near the top. The cylinder can be spun to allow one object from the bottom set to be positioned just behind the bottom opening where participants can reach in to grasp an object hidden from view by two broom brushes. As well, the top set can be spun separately from the bottom to allow one object to be positioned in a 6.5-cm × 10-cm window allowing the participant to see one object at a time. The top set could also be spun so that no object was visible (for participants in the haptic learning condition).

Each object was paired with a name from a list of eight non-word labels. The pairings of non-word names and objects was changed between participants, ensuring that a relation between name and object did not confound learning measures at the group-average level. The non-word labels were four-

letter non-words that are easily pronounced but do not have any semantic meaning: yoot, baiv, fint, malg, verp, grov, demb, and jorl.

The experiment was run using Superlab 4.5, and later Superlab 5.0, as well as a Cedrus SV1 Voice Key to trigger experimental sequences. Participants wore a microphone-equipped headset to receive auditory cues and trigger voice-onset events.

Procedure

The experiment comprised a learning phase and an experimental phase. The timeline of experimental events is depicted in Fig. 2.

Learning phase: Participants were randomly assigned to one of three learning conditions—haptic (22 participants), visual (23 participants), or bimodal (25 participants)—and were presented with sequences of learning and test trials.

Haptic learning condition: Participants sat at a table in front of the apparatus containing the novel objects, wearing headphones, with their hand resting on the table. Participants were instructed to keep their eyes closed while the experimenter set up each trial by rotating the object (s) in the correct position. The auditory cue “reach” signaled the beginning of learning trials. Participants reached in the apparatus with their right hand to grasp the object in the bottom opening. Participants were given 5 s to grasp the object while the experimenter said the name of the object. The auditory cue “end of trial” signaled the end of trials, at which point participants removed their hand. Participants were presented with each of the eight objects once in random order.

Once participants completed eight learning trials, they completed 16 test trials: one set of eight visual identification trials (one for each object), and one set of eight haptic identification trials (one for each object). The order of the two testing sets (haptic and visual) was reversed between each block of eight learning and 16 test trials. At the beginning of test trials participants were asked to keep their eyes closed while the experimenter rotated the objects in the correct position. The auditory cue ‘open’ (for a visual identification trial) or ‘reach’ (for a haptic identification trial) signaled the beginning of a test trial. At this point, participants would either open their eyes to see an object through the top window, or reach with their hand and grasp an object through the bottom window—only one object was presented during each trial. A card bearing the eight non-words in pseudo-random order was also placed in front of participants to ensure mistakes would reflect confusion between objects and not an inability to remember a specific non-word (a new card with a different pseudo-random order was placed at the beginning of each testing block). Participants were instructed to name each object, and make their best guess if unsure of the answer; no feedback was provided. Participants’ verbal responses triggered the “end

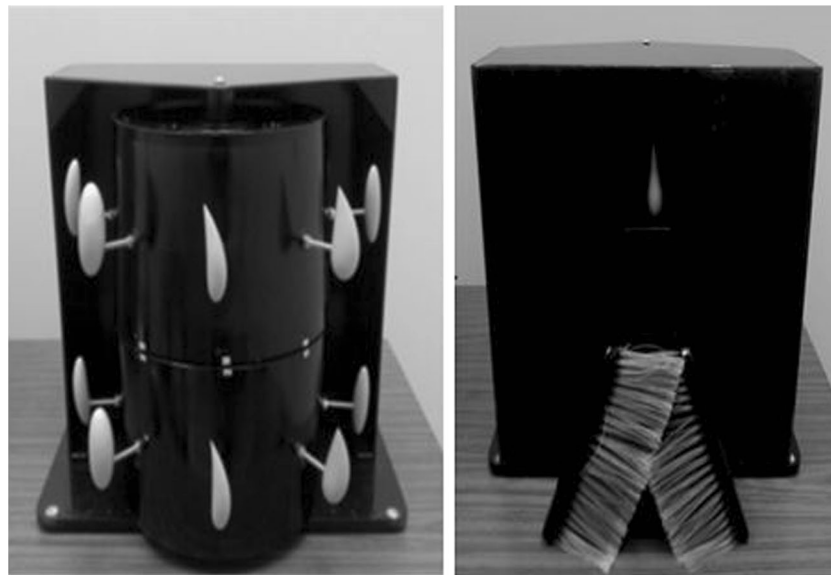


Fig. 1 Rotating apparatus used to display objects visually and/or haptically

of trial” cue, at which point participants would remove their hand or close their eyes. The experimenter then recorded participants’ response using the keyboard and rotate the objects in place for the next test trial. All eight objects were presented in random order once for visual identification, and once for haptic identification for a total of 16 test trials. Once participants named all objects once visually and once haptically, they were

presented with eight more learning trials followed by 16 more test trials. This sequence of interleaved learning and test trials continued until participants had correctly identified all objects visually and haptically twice in a row, reaching criterion. At this point the learning phase was discontinued.

Visual learning condition: The sequence of events for participants in the visual learning condition was identical to that

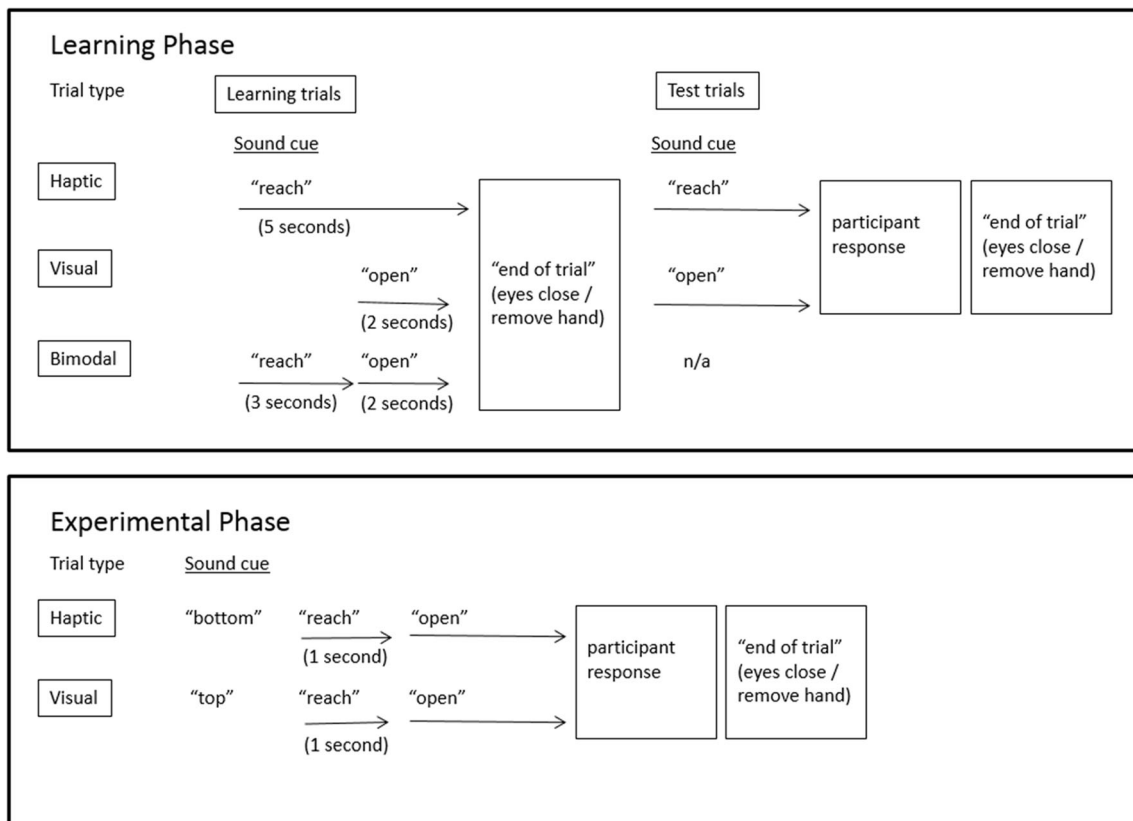


Fig. 2 Timeline of events in Experiment 1

of participants in the haptic learning condition, except for one main difference: during learning trials, instead of reaching and grasping an object, participants heard the auditory cue “open,” after which they opened their eyes and saw an object through the top opening of the apparatus. Participants were given 2 s to look at each object while the experimenter said the name of the object before the auditory cue “end of trial” indicated that participants should close their eyes.

Bimodal learning condition: The sequence of events for participants in the bimodal learning condition was identical to that of other participants, except that during learning trials, participants both viewed and grasped objects. Participants first heard the command “reach,” followed by the command “open” 3 s later—this allowed them to feel an object with their right hand for 5 s AND concurrently see an identical object for the last 2 s (also while the experimenter said the name of the object).

The 2-s (for visual exploration) and 5-s (for haptic exploration) durations were selected to compensate for the time to move the hand and to be consistent with other visuo-haptic shape perception experiments (see Gallace & Spence, 2009; Lacey & Campbell, 2006; Lacey, Pappas, Kreps, Lee, & Sathian, 2009; Peltier, Stilla, Mariola, LaConte, Hu, & Sathian, 2007).

Experimental phase: All participants completed an identical experimental phase. Participants were first presented with 16 reminder trials that were identical to the learning trials of the learning phase. Each object was presented once visually and once haptically in random order—there were no bimodal reminder trials. Participants then completed a speeded identification task. Participants kept their eyes closed while the experimenter rotated the objects in the correct position. The experimenter then signaled whether participants should identify the visually presented object by saying “top” or the haptically presented object by saying “bottom.” The auditory cue “reach” signaled that participants should reach with their right hand into the bottom opening to grasp the bottom object. One second later, the auditory cue “open” signaled that they could open their eyes to look at the visually displayed object; this cue served as the recording onset for participants’ reaction time. The 1-s delay ensured that participants had time to reach and grasp the object before opening their eyes. Participants then ‘named’ the target object as fast as possible, without making too many mistakes, and the onset of their vocal response was recorded as the reaction time. No feedback was provided, and the experimenter then recorded the participant’s answer with a key press. Participants completed 224 trials in four blocks of 56 trials, with short participant-paced breaks in between. Half of the trials required participants to identify the haptically presented object at the bottom and half the trials required participants to identify the visually presented object at the top. Crucially, half of the 112 haptic and the 112 visual trials were congruent (two copies of the same object were

presented), while the other half of the trials were incongruent (the visual and the haptic objects were two different objects). For the 56 incongruent trials in a given modality, each of the eight objects was presented seven times for identification, once with each of the seven other possible objects. For the 56 congruent trials in a given modality, each of the eight objects was presented seven times for identification—each time with another copy of itself in the second modality. Altogether, participants completed 112 haptic identification trials, composed of 56 congruent trials (seven trials for each of the eight objects) and 56 incongruent trials (seven trials for each of the eight objects), and 112 visual identification trials, composed of 56 congruent trials (seven trials for each of the eight objects) and 56 incongruent trials (seven trials for each of the eight objects). The order of individual trials was randomized, and participants took approximately 2 h to complete the experiment.

Results

Learning phase

First, we recorded the number of blocks needed to reach criterion for visual and haptic identification for participants in each learning condition. Data were entered in a 2 (testing modality) \times 3 (learning condition) mixed ANOVA where the repeated factor was testing modality (visual or haptic) and the between factor was learning condition (visual, haptic, or bimodal). The analysis revealed a main effect of testing modality, $F(1, 67) = 15.329$, $p = 0.001$, $partial \eta^2 = 0.155$. Generally, participants needed more blocks to reach criterion when identifying novel objects haptically (mean = 9.0 blocks) than when identifying objects visually (mean = 8.3 blocks). This main effect was qualified by a two-way interaction between learning condition and testing modality, $F(1, 67) = 7.572$, $p = 0.004$, $partial \eta^2 = 0.154$. Planned directional matched t -tests confirmed that participants in the visual learning condition required fewer blocks to reach criterion when asked to identify novel objects visually (mean = 8.3 blocks) than when asked to identify novel objects haptically (mean = 9.6 blocks), $t(22) = 3.533$, $p = 0.001$, Cohen’s $d = 0.546$. Similarly, participants in the bimodal learning condition required fewer blocks to reach criterion when asked to identify novel objects visually (mean = 7.8 blocks) than when asked to identify novel objects haptically (mean = 8.8 blocks), $t(24) = 2.739$, $p = 0.011$, Cohen’s $d = 0.418$. However, and contrary to our prediction, participants in the haptic learning condition required similar numbers of blocks to reach criterion when asked to identify novel objects visually (mean = 8.9 blocks) and when asked to identify novel objects haptically (mean = 8.6 blocks), $t(22) = -1.240$, not significant (ns). The results are depicted in Fig. 3.

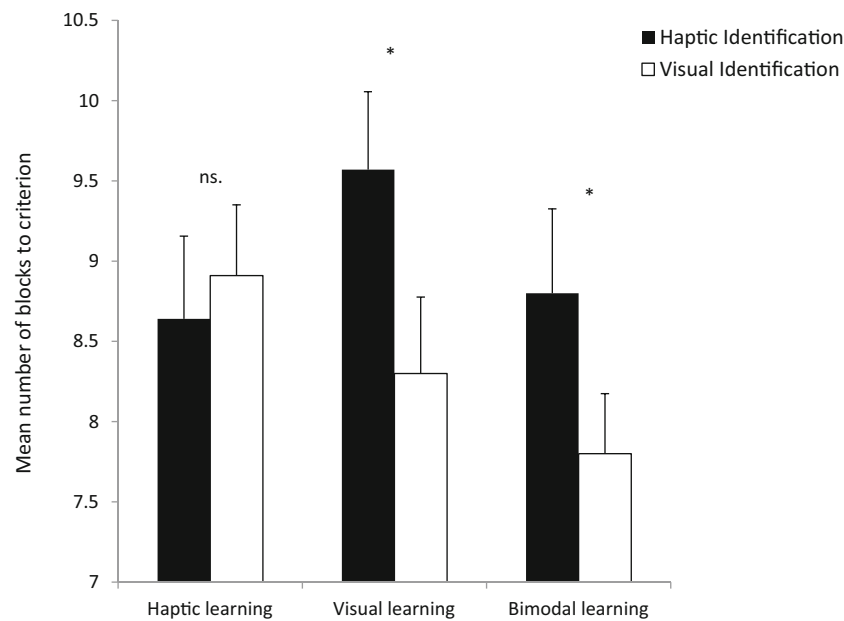


Fig. 3 Mean number of blocks to criterion and standard error for the visual and haptic identification of novel objects, for participants in the haptic, visual, and bimodal learning conditions. Asterisks indicate

significant differences in mean number of blocks to criterion between haptic identification and visual identification

Next, we assessed the patterns of confusion in memory. We classified each error according to which other object the presented object was confused with. For example, if object 0,0,1 was presented but participants used the label associated with object 0,1,1, the error was classified as a confusion between objects that are 1 CB apart. If object 0,0,1 was presented but participants used the label associated with object 1,1,1, the error was classified as a confusion between objects that are 2 CBs apart. Because there were 12 pairs of objects that were 1 CB apart, 12 pairs of objects that were 2 CBs apart, and 4 pairs of objects that were 3 CBs apart, the number of confusion errors in each category was divided by the number of pairs of objects that could generate such errors. These averages were entered in a 2 (testing modality) \times 3 (learning condition) \times 3 (CB distance) mixed ANOVA for which the repeated factors were testing modality (visual or haptic) and CB distance (1 CB, 2 CBs, or 3 CBs), and the between factor was learning condition (visual, haptic, or bimodal). The analysis revealed a main effect of CB distance, $F(2,134) = 23.644$, $p < 0.001$, $partial \eta^2 = 0.691$. Generally, novel objects that were 1 CB apart (mean = 1.37 errors) were confused more often than objects that were 2 CBs apart (mean = 0.69 errors), $t(69) = 13.143$, $p < 0.001$, $Cohen's d = 1.557$, and more often than objects that were 3 CBs apart (mean = 0.618 errors), $t(69) = 12.934$, $p < 0.001$, $Cohen's d = 1.621$. Objects that were 2 CBs apart were also confused more often than objects that were 3 CBs apart, $t(69) = 1.827$, $p = 0.036$, $Cohen's d = 0.172$. This main effect was qualified by a two-way interaction between learning condition and CB distance, $F(4, 134) = 3.262$, $p = 0.014$, $partial \eta^2 = 0.089$. There was also an interaction

between learning condition and testing modality, $F(2, 67) = 4.779$, $p = 0.011$, $partial \eta^2 = 0.125$, as well as a three-way interaction between all factors, $F(4,134) = 3.315$, $p = 0.013$, $partial \eta^2 = 0.09$. The results are presented in Fig. 4.

Recall that we hypothesized that performance would be best when testing modality matched learning conditions (*see* Easton, Greene, & Srinivas, 1997; Hershberger & Misceo, 1996; Newell et al., 2005). We therefore analyzed the interaction between learning condition and testing modality using planned repeated measures t -tests. The results paralleled those observed for blocks to criterion. Participants in the visual learning condition produced more errors when identifying objects haptically (mean = 0.96 errors) than when identifying objects visually (mean = 0.88 errors), $t(22) = 1.869$, $p = 0.035$, $Cohen's d = 0.193$. Participants in the bimodal learning condition also produced more errors when identifying objects haptically (mean = 0.92 errors) than when identifying objects visually (mean = 0.81 errors), $t(24) = 2.496$, $p = 0.01$, $Cohen's d = 0.275$. Finally, participants in the haptic learning condition did not produce more errors when identifying objects visually (mean = 0.91 errors) than when identifying objects haptically (mean = 0.85 errors), $t(21) = 1.665$, ns .

To uncover the source of the three-way interaction, we ran individual repeated measures ANOVA for which the factors were learning condition (haptic, visual, or bimodal) and testing modality (visual or haptic)—one analysis was run for the data at each CB distance. The analysis of errors that were 1 CB apart revealed a main effect of testing modality, $F(1, 67) = 3.949$, $p = 0.051$, $partial \eta^2 = 0.056$, which was qualified by a two-way interaction between learning condition and testing

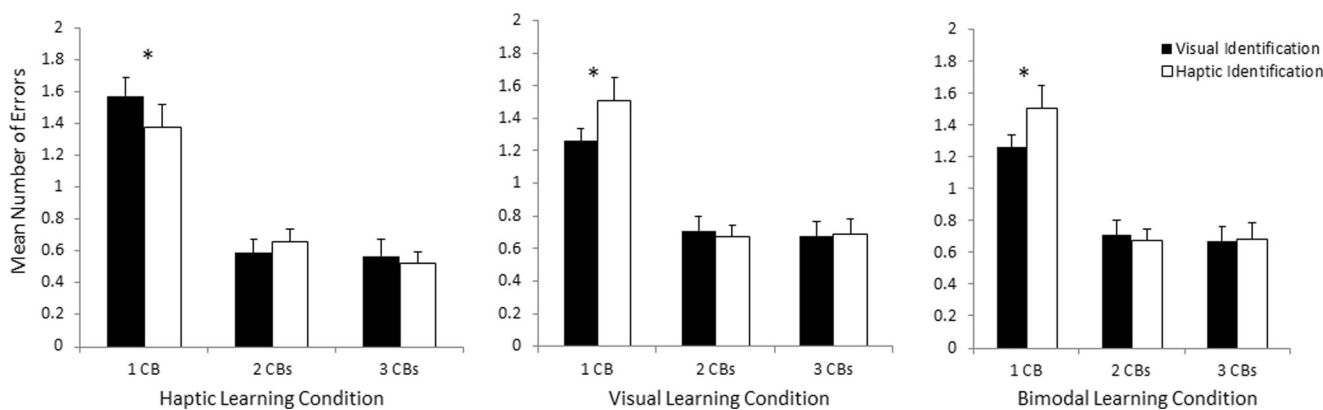


Fig. 4 Mean number of errors between objects that are 1 CB, 2 CBs, and 3 CBs apart (and standard error) for the visual and haptic identification of novel objects, for participants in the haptic learning condition, visual

learning condition, and bimodal learning conditions. Asterisks indicate significant differences in mean number of errors between haptic identification and visual identification

modality, $F(1,67) = 8.746$, $p < 0.001$, $partial \eta^2 = 0.207$. Bonferroni-corrected t -tests revealed that for participants in the haptic learning condition, there were more errors for visual identification (mean = 1.58 errors) than for haptic identification (mean = 1.38 errors), $t(21) = -2.666$, $p < 0.01$, Cohen's $d = 0.303$. In contrast, for participants in the visual learning condition, there were more errors for haptic identification (mean = 1.51 errors) than for visual identification (mean = 1.261 errors), $t(22) = 2.283$, $p = 0.016$, Cohen's $d = 0.42$. This also occurred for participants in the bimodal learning condition, there were more errors for haptic identification (mean = 1.347 errors) than for visual identification (mean = 1.103 errors), $t(24) = 3.524$, $p = 0.001$, Cohen's $d = 0.576$. The analyses of errors that were 2 CBs and 3 CBs apart revealed no significant effects.

Recall that one of the limitations of using real objects in identification studies is that real objects do not offer the possibility to know which object dimensions are used to make comparisons or identifications. In the previous analysis, all dimensions were grouped together for comparison purposes. In a separate analysis, we used only confusions between objects that were 1 CB apart—these are the comparisons for which we know which dimension was confused. We calculated the proportion of 1 CB errors that were based on curvature, tapering, and thickness, and entered the inverse sine function (\arcsin) of the square root these proportions in a 3 (learning condition) \times 2 (testing modality) \times 3 (object dimension) mixed design ANOVA, where the between subject factor was learning condition, and the repeated measures were testing modality and object dimension. The analysis revealed a main effect of object dimension, $F(2, 134) = 36.176$, $p < 0.001$, $partial \eta^2 = 0.351$, which was qualified by a two-way interaction between object dimension and testing modality, $F(2, 134) = 3.723$, $p = 0.027$, $partial \eta^2 = 0.053$. No other effects were significant; the results are presented in Fig. 5. Bonferroni-corrected t -tests revealed that when participants identified objects haptically, they confused curved and

uncurved objects (mean = 0.797) more often than tapered and untapered objects (mean = 0.493), $t(69) = 6.777$, $p < 0.001$, Cohen's $d = 0.649$, or than thick and thin objects (mean = 0.489), $t(69) = 7.692$, $p < 0.001$, Cohen's $d = 0.771$. However, there were as many confusion errors between tapered and untapered objects as there were between thick and thin objects, $t(69) = -0.126$, ns . When participants identified objects visually, they confused curved and uncurved objects (mean = 0.730) more often than tapered and untapered objects (mean = 0.493), $t(69) = 5.280$, $p < 0.001$, Cohen's $d = 0.667$, or than thick and thin objects (mean = 0.556), $t(69) = 4.044$, $p < 0.001$, Cohen's $d = 0.461$. Again, there were as many confusion errors between tapered and untapered objects as there were between thick and thin objects, $t(69) = -1.212$, ns . Essentially, the same pattern of confusion errors occurred in each testing modality where there were more errors along the curvature dimension than along the tapering or thickness dimension, but the differences observed were larger for haptic identification than for visual identification.

We also calculated correlations for each participant's overall confusion matrix to evaluate whether confusion errors in haptic identification corresponded to confusion errors in visual identification. Each participant's correlation was transformed using the Fisher Z transform, and we used t -tests to evaluate whether the group-average correlation was significantly different from 0 in each of the learning conditions. All group-average correlations were significantly greater than 0. For participants in the haptic learning condition, the mean correlation (0.434) was significantly greater than 0, $t(21) = 6.511$, $p < 0.001$, Cohen's $d = 2.842$. For participants in the visual learning condition, the mean correlation (0.306) was significantly greater than 0, $t(22) = 5.986$, $p < 0.001$, Cohen's $d = 2.552$. For participants in the bimodal learning condition, the mean correlation (0.351) was significantly greater than 0, $t(24) = 7.935$, $p < 0.001$, Cohen's $d = 3.239$. Finally, we entered the individual correlations into a univariate analysis for which the between factor was learning condition.

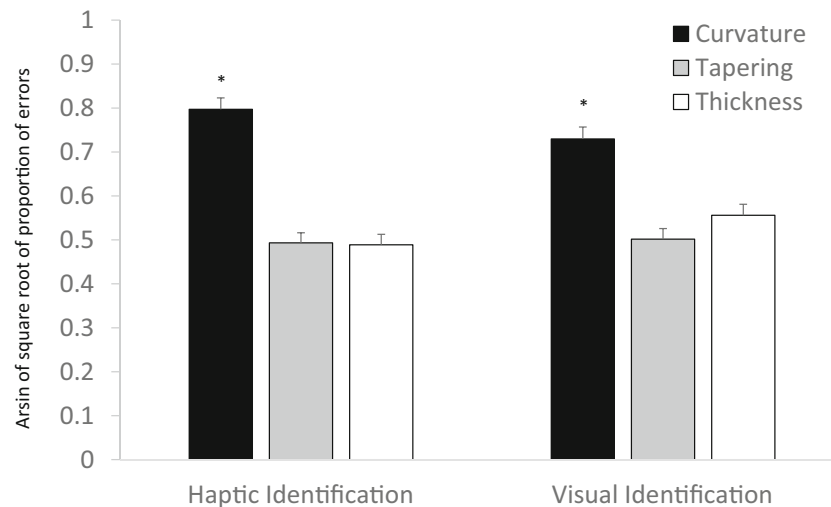


Fig. 5 Mean arsin of square root of proportion of errors (standard error) for confusion errors along the dimension of curvature, tapering, and thickness for haptic and visual identification. Asterisks indicate

significant differences in mean proportion of errors for novel objects that differ on curvature vs novel objects that differ on another dimension

The analysis revealed that learning condition did not influence the similarity of haptic and visual identification, $F(2, 67) = 1.393$, *ns*.

Finally, we collected reaction times for the last two blocks of learning trials; the onset of reaction time recording coincided with the auditory cue to reach or open their eyes, and was terminated when participants initiated their verbal response. We limited the analysis to the two blocks where performance accuracy was stable (all responses were correct). We averaged reaction time for haptic identification and visual identification, and entered the data in a 3 (learning condition) \times 2 (testing modality) mixed design ANOVA, for which the between factor was learning condition (haptic, visual, or bimodal), and the repeated factor was testing modality (haptic identification, visual identification). The analysis revealed a main effect of learning condition $F(2, 67) = 26.513$, $p < 0.001$, *partial* $\eta^2 = 0.442$ a main effect of testing modality, $F(1, 67) = 88.163$, $p < 0.001$, *partial* $\eta^2 = 0.568$, as well as an interaction between the two factors, $F(2, 67) = 13.912$, $p < 0.001$, *partial* $\eta^2 = 0.293$. The main effect of testing modality reflects that visual identification (mean = 3,298 ms) was consistently faster than haptic identification (mean = 5,850 ms). Because the longer reaction times to identify objects haptically may reflect the time it takes to reach and grasp each object, we analyzed the interaction by contrasting performance within each testing modality. Bonferroni corrected *t*-tests showed that individual groups differed in how quickly they could identify objects haptically or visually. Participants in the haptic learning condition identified objects haptically (mean = 4,438 ms) and visually (mean = 2,628 ms) as quickly as participants in the visual learning condition (mean = 4,382 ms for haptic identification and 3,130 ms for visual identification), $t(43) = -0.098$, *ns* for haptic identification, and $t(43) = 1.444$, *ns* for visual identification, respectively. However, participants in the

bimodal learning condition were consistently slower (mean = 8,444 ms for haptic identification, and 4,045 ms for visual identification) than participants in the haptic learning condition [$t(45) = -5.656$, $p < 0.001$, Cohen's $d = 1.653$ for haptic identification, and $t(45) = -5.023$, $p < 0.001$, Cohen's $d = 1.468$ for visual identification], and than participants in the visual learning condition [$t(46) = -5.480$, $p < 0.001$, Cohen's $d = 1.583$ for haptic identification, and $t(46) = -3.009$, $p < 0.005$, Cohen's $d = 0.666$ for visual identification]. The results are presented in Fig. 6.

Experimental phase

Learning condition, testing modality and congruence

Reaction times for correct trials were trimmed recursively at three standard deviations (*see* Van Selst & Jolicoeur, 1994—this resulted in the removal of 2% of the raw data), averaged, and entered in a 3 (learning condition) \times 2 (testing modality) \times 2 (congruence) mixed ANOVA for which the between factor was learning condition (haptic, visual, or bimodal, as per the learning phase) and the repeated factors were testing modality (haptic or visual) and congruence (congruent or incongruent). The analysis revealed a main effect of testing modality, $F(1, 67) = 19.175$, $p < 0.001$, *partial* $\eta^2 = 0.223$; generally, haptic identification (mean = 2,668 ms) was slower than visual identification (mean = 2,494 ms). There was also a main effect of congruence $F(1, 67) = 29.086$, $p < 0.001$, *partial* $\eta^2 = 0.303$; generally, reaction time was faster for congruent trials (mean = 2,513 ms) than for incongruent trials (mean = 2,648 ms). These main effects were qualified by two two-way interactions. First, there was an interaction between testing modality and congruence, $F(2, 67) = 6.526$, $p = 0.013$, *partial* $\eta^2 = 0.089$. Incongruent information produced

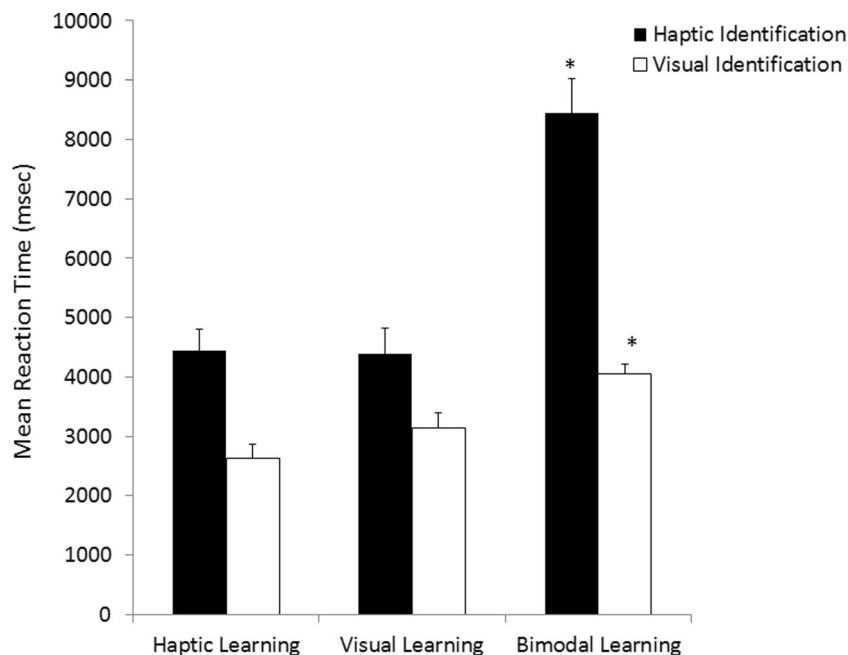


Fig. 6 Mean reaction time (standard errors) for the correct identification of haptic and visual stimuli for participants in the haptic, visual, and bimodal learning conditions. Asterisks indicate significant differences in

reaction time for participants in the bimodal learning condition vs participants in other learning conditions

interference for both haptic and visual identification. However, the interference observed during haptic identification (2,626 ms for congruent trials versus 2,709 ms for incongruent trials), $t(69) = -2.276, p = 0.013$, Cohen’s $d = 0.100$ generated a smaller effect size than the interference observed during visual identification (2,400 ms for congruent trials versus 2,587 ms for incongruent trials), $t(69) = -7.333, p < 0.001$, Cohen’s $d = 0.266$ (see Fig. 7). There was also an interaction between learning condition and testing modality, $F(2,67) = 3.125, p = 0.05, partial \eta^2 = 0.285$. Planned paired samples t -tests revealed that, similar to what was observed in the learning phase, for participants in the visual learning condition, visual identification (mean = 2,447 ms) was faster than haptic identification (mean = 2,760 ms), $t(22) = 4.496, p > 0.001$, Cohen’s $d = 0.341$. For participants in the bimodal learning condition, visual identification (mean = 2,521 ms) was also faster than haptic identification (mean = 2,655 ms), $t(23) = 1.888, p = 0.035$, Cohen’s $d = 0.165$. In contrast, for participants in the haptic learning condition, reaction time for visual identification (mean = 2,513 ms) did not significantly differ from reaction time for haptic identification (mean = 2,589 ms), $t(21) = 1.178, ns$ (see Fig. 8). No other effects reached significance.

A second 3 (learning condition) \times 2 (testing modality) \times 2 (congruence) mixed ANOVA was performed on the arcsin of the square root of the proportion of errors committed in each condition. Again, the between factor was learning condition

(haptic, visual, or bimodal) and the repeated measures factors were testing modality (haptic identification or visual identification) and congruence (congruent or incongruent). The analysis revealed three main effects. There was a main effect of congruence, $F(1,67) = 30.380, p < 0.001, partial \eta^2 = 0.312$.

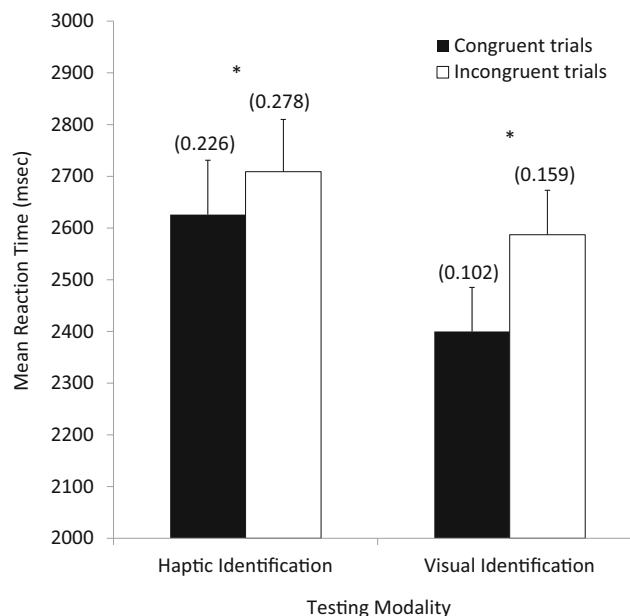


Fig. 7 Mean reaction time (arcsin of square root of proportion of errors) and standard errors for the haptic and visual object identification during congruent and incongruent trials. Asterisks denote significant differences in reaction time between congruent and incongruent trials

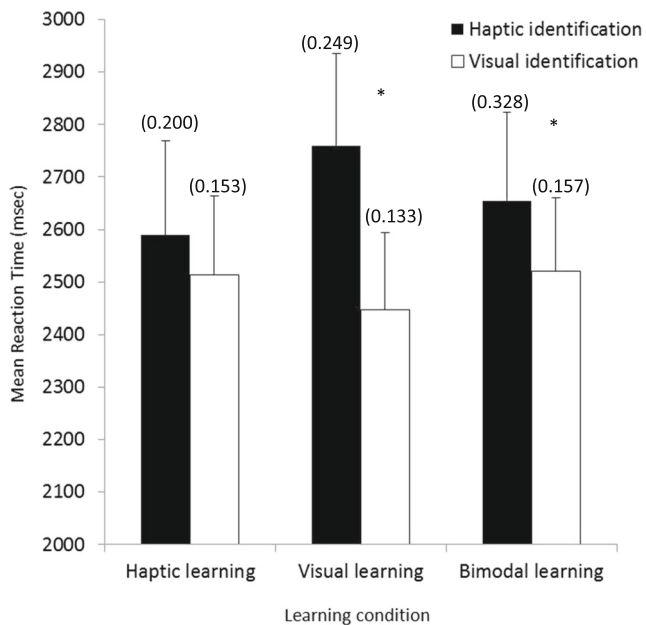


Fig. 8 Mean reaction time (arsin of square root of mean proportion of errors) and standard errors for the haptic and visual identification of objects for participants in the haptic learning condition, the visual learning condition, and the bimodal learning condition. Asterisks indicate significant differences in reaction time and error rates between haptic identification and visual identification

Generally, participants were more accurate during congruent trials (mean = 0.164) than during incongruent trials (mean = 0.219). There was also a main effect of learning condition, $F(2,67) = 4.344$, $p = 0.017$, $partial \eta^2 = 0.115$. Generally, participants in the visual learning condition (mean = 0.206) were as accurate as participants in the haptic learning condition (mean = 0.186), $t(43) = 0.365$, *ns*. Interestingly, participants in the bimodal learning condition (mean = 0.263) were *less accurate* than participants in the visual learning condition, $t(46) = 2.301$, $p = 0.026$, Cohen's $d = 0.665$, or than participants in the haptic learning condition, $t(45) = 3.093$, $p = 0.003$, Cohen's $d = 0.904$. There was also a main effect of testing modality, $F(1,67) = 78.952$, $p < 0.001$, $partial \eta^2 = 0.541$. These main effects were qualified by a two-way interaction between learning condition and testing modality, $F(2,67) = 5.324$, $p = 0.007$, $partial \eta^2 = 0.137$. The pattern of error rates paralleled the patterns of reaction times. Planned directional paired-sample *t*-tests revealed that participants in the visual learning condition and the bimodal learning condition were more accurate when identifying objects visually (means = 0.133 and 0.157, respectively) than when identifying objects haptically (means = 0.249 and 0.328, respectively), $t(22) = 6.124$, $p < 0.001$, Cohen's $d = 1.249$ for participants in the visual learning condition, and $t(24) = 6.370$, $p < 0.001$, Cohen's $d = 1.402$. Participants in the haptic learning condition were equally accurate when identifying objects visually (mean = 0.153) and haptically (mean = 0.200), $t(21) = 1.998$, *ns*.

Object similarity

We carried out a second analysis to test the hypothesis that interference was mediated by similarity. We only used incongruent trials and calculated mean reaction times for trials where the distracter object was 1 CB apart, 2 CBs apart, and 3 CBs apart. The data were entered in a 3 (learning condition) \times 2 (testing modality) \times 3 (CB distance) mixed ANOVA, where the between factor was learning condition (haptic, visual, or bimodal) and the repeated factors were testing modality (haptic identification, or visual identification) and CB distance (1 CB, 2 CBs, or 3 CBs). The analysis only revealed a main effect of testing modality, $F(1, 67) = 5.183$, $p = 0.026$, $partial \eta^2 = 0.072$. During these incongruent trials, visual identification (mean = 2,589 ms) was faster than haptic identification (mean = 2,699 ms). No other effects were significant.

We also examined the effect of object similarity on accuracy by analyzing the arsin of the square root of the proportion of errors for trials where the distracter object was 1 CB apart, 2 CBs apart, and 3 CBs apart. The data was first entered in a 3 (learning condition) \times 2 (testing modality) \times 3 (CB distance) mixed ANOVA, where the between factor was learning condition (haptic, visual, or bimodal) and the repeated factors were testing modality (haptic identification or visual identification) and CB distance (1 CB, 2 CBs, or 3 CBs). The analysis revealed two main effects: a main effect of learning condition, $F(2, 67) = 5.495$, $p = 0.006$, $partial \eta^2 = 0.141$, and a main effect of testing modality, $F(1, 67) = 38.855$, $p < 0.001$, $partial \eta^2 = 0.367$. Overall, visual identification (mean = 0.120) was more accurate than haptic identification (mean = 0.242). Also, participants in the visual learning condition (mean = 0.157) and in the haptic learning condition (mean = 0.150) were equally accurate, $t(43) = 0.812$, *ns*. However, as was observed when all trials (congruent and incongruent) were analyzed, participants in the bimodal learning condition (mean = 0.235) were less accurate than participants in the visual learning condition $t(46) = 2.714$, $p = 0.009$, Cohen's $d = 0.784$, and less accurate than participants in the haptic learning condition, $t(45) = 2.859$, $p = 0.006$, Cohen's $d = 0.836$. These main effects were qualified by a two-way interaction between both factors, $F(2, 67) = 6.228$, $p = 0.003$, $partial \eta^2 = 0.157$ (see Fig. 9). This interaction paralleled that observed in the learning phase. Planned repeated-measures *t*-tests showed that for participants in the haptic learning condition, haptic identification (mean = 0.219) was as accurate as visual identification (mean = 0.168), $t(21) = 1.453$, *ns*. For participants in the visual learning condition, haptic identification (mean = 0.254) was less accurate than visual identification (mean = 0.144), $t(22) = 5.013$, $p < 0.001$, Cohen's $d = 1.033$. Similarly for participants in the bimodal learning condition, haptic identification (mean = 0.360) was less accurate than visual identification (mean = 0.165), $t(24) = 5.733$, $p < 0.001$, Cohen's $d = 1.42$.

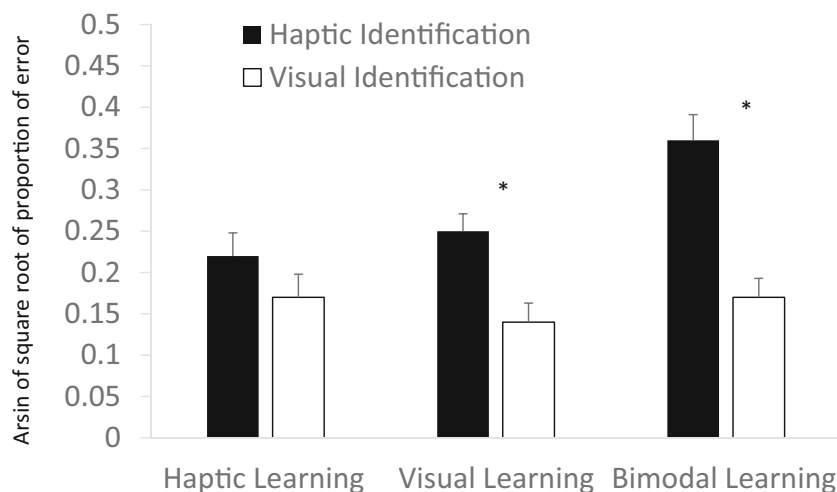


Fig. 9 Mean arsin of square root of proportion of errors and standard error for haptic identification and visual identification for participants in the haptic, visual, and bimodal learning conditions. Asterisks denote significant differences between haptic and visual identification

Discussion

Experiment 1 demonstrated that haptic and visual identification are highly similar, and that matching learning and testing conditions does not necessarily lead to better performance. Consistent with our hypothesis, the error patterns were similar across testing modalities: in both testing modalities, similar objects were confused more often than distinct objects. Also, curvature was the source of confusion errors more often than tapering or thickness, independently of testing modality. Finally, the pattern of haptic identification errors correlated with the pattern of visual identification errors independently of learning condition: individual participants confused the same objects during haptic and visual identification. Our results demonstrate that haptic and visual identification follow similar principles and are therefore comparable, at least when identification is based on shape information.

However, contrary to our hypothesis, identification performance was not always best when testing modality matched learning condition. If the learning condition included a visual component, participants performed better during visual identification: participants in the visual and bimodal learning conditions performed better (i.e., they were faster and more accurate, during both the learning phase and the experimental phase) during visual identification than during haptic identification. This was not true for participants in the haptic learning condition for which there was no difference in performance: they were equally fast and accurate during haptic and visual identification.

Finally, contrary to our hypothesis, object similarity did not impact performance in the experimental phase. This may be the result of the two types of identification trials that participants completed during the learning phase: though participants were exposed to haptic, visual, or bimodal learning

trials, during test trials all participants were asked to identify objects visually and haptically. This was done to ensure participants could identify objects in both modalities—if we had failed to detect interference, this could have been due to a lack of familiarity with the untested modality. However, it is possible that the absence of an effect of similarity is due to the blurring of the learning conditions.

We therefore asked a new group of participants to complete only the multisensory experimental phase. However, participants may have much difficulty if asked to immediately identify eight novel objects without repeated exposure—it did take on average more than seven blocks (and therefore seven individual exposures to objects) for participants to reach criterion in Experiment 1. Consequently, instead of asking participants to identify objects, we asked participants to indicate whether the two presented objects were the same or different. We hypothesized that reaction times to incongruent trials would be affected by object similarity: we specifically expected slower reaction times when similar objects were presented, and faster reaction times when distinct objects were presented.

Experiment 2

Method

Participants

Participants were 21 undergraduate university students from an introductory psychology course at Mount Allison University (12 females, mean age = 18.57 years). All participants reported being right-handed and having normal or corrected-to-normal vision. Participants received 0.5 credits for each half hour of participation in the study.

Materials

The materials were the same object and rotating apparatus used in Experiment 1.

Procedure

The procedure was identical to the experimental phase of Experiment 1, except for one key difference: participants were instructed to indicate whether the two objects were identical or different by pressing one of two keys. Half of the participants were instructed to press the “same” key with their index finger and the “different” key with their middle finger, while the other half of the participants did the opposite. Reaction time was recorded at the button press. No feedback was provided, and the study took approximately 1 h to complete.

Results

Reaction times to correct trials were trimmed recursively at three standard deviations (this resulted in the removal of 3.3% of the data). Reaction times to incongruent trials were entered in a one-way repeated measures ANOVA, for which the factor was CB Distance (1 CB, 2 CBs, 3 CBs). The main effect of CB Distance was significant, $F(1, 20) = 12.896$, $p = 0.002$, $\text{partial } \eta^2 = 0.39$. Planned repeated measures t -tests revealed that participants' reaction times were slower when presented with two objects that were more similar than when presented with objects that were more distinct. Specifically, participants responded ‘different’ more slowly when objects were 1CB apart (mean = 2,628 ms) than when objects were 2 CBs apart (2,138 ms), $t(20) = 3.863$, $p = 0.001$; Cohen's $d = 0.331$, and than when objects were 3 CBs (mean = 1,997 ms) apart, $t(20) = 3.591$, $p = 0.002$; Cohen's $d = 0.576$. Furthermore, participants responded ‘different’ more slowly when objects were 2 CBs apart than when objects were 3 CBs apart, $t(20) = 2.099$, $p = 0.049$; Cohen's $d = 0.26$.

The arcsin of the square root of the mean proportion of errors for incongruent trials were also entered in a repeated measures ANOVA that confirmed a main effect of CB Distance, $F(1, 20) = 61.252$, $p < 0.001$, $\text{partial } \eta^2 = 0.754$. Planned repeated t -tests showed that there were more errors when objects differed by 1 CB (mean = 0.374) than when objects differed by 2 CBs (mean = 0.098), $t(20) = 9.722$, $p < 0.001$; Cohen's $d = 1.888$, or than when objects differed by 3 CBs (mean = 0.042), $t(20) = 8.958$, $p < 0.001$; Cohen's $d = 2.359$. Participants also made significantly more errors when objects differed by 2 CBs than when objects differed by 3 CBs $t(20) = 1.838$, $p < 0.05$, Cohen's $d = 0.393$.

Discussion

Consistent with our hypothesis, object similarity impacted participants' ability to evaluate whether two objects are the same or are different. Participants were slower and less accurate when objects were similar than when objects were distinct. Object similarity can clearly influence performance when objects are novel.

Why did object similarity not impact reaction time when participants were asked to identify objects after a certain level of performance was obtained? During the test trials of the learning phase of Experiment 1, participants were asked to identify objects *both* visually and haptically. It is possible that this created more extensive representations in memory than the three coordinate system we used to describe each object. Instead, the extensive practice (i.e., haptic + visual identification) could have led to a six-coordinate system (three visual dimensions and three haptic dimensions), creating greater distance in object space than anticipated. In the original studies of Desmarais et al. (2007), the more distinct objects (2 CBS and 3 CBs apart) were grouped together and compared to the more similar objects (1 CB apart) because error rates were equivalent at greater distances in object space. Asking participants to practice haptic identification and visual identification may have increased the distance in object space, making even the closest objects (1 CB) seem more distinct (they actually differ on 1 visual CB and 1 haptic CB) and preventing us from detecting an impact of object similarity in Experiment 2.

In order to test this possibility, we asked a new group of participants to complete the same procedure outlined in Experiment 1 with one important change: during the learning phase, participants were only exposed to objects in one modality. That is, participants explored and identified objects only haptically or only visually. We expected that this change in procedure would reduce the potential distance in object space and allow us to measure an impact of object similarity on reaction times and error rates when two different objects were presented for identification. We expected that participants would identify objects more slowly and less accurately when the distracter object was similar (1 CB) to the target object than when the distracter object was distinct (2 or 3 CBs) from the target object.

Experiment 3

Method

Participants

Participants were 40 undergraduate university students (20 females, mean age = 19.77 years) from an introductory psychology course at Mount Allison University. All participants

reported being right-handed and having normal or corrected-to-normal vision, and they received 0.5 credits for each half hour they participated in the study.

Materials

The materials were the same objects, rotating apparatus, and nonword names as those used in Experiment 1.

Procedure

The procedure was nearly identical to that of Experiment 1, except that participants only experienced the objects in one modality (haptic *or* visual) during the test trials of the learning phase. Twenty participants were presented with sequences of eight haptic learning trials followed by eight haptic test trials, while another 20 participants were presented with sequences of eight visual learning trials followed by eight visual test trials (we did not run participants in the bimodal condition). These participants then completed an experimental phase identical to that described in Experiment 1.

Results

Learning phase

We compared the number of blocks it took participants in the visual and haptic learning conditions to reach criterion and entered this data in an independent samples *t*-test. There was no difference in number of blocks to criterion for participants in the visual (mean = 9.55 blocks) and haptic (mean = 10.25 blocks) learning condition, $t(38) = -0.725$, *ns*.

We then assessed the patterns of confusion in memory. As in Experiment 1, each error was classified according to which other object the presented object was confused with, and the number of confusion errors in each category was divided by the number of pairs of objects that could generate such errors. The mean number of errors were entered in a 2 (learning condition) \times 3 (objects similarity) mixed design ANOVA, where the between-subject factor was learning condition and the repeated factor was object similarity. There was a main effect of object similarity, $F(2,76) = 54.76$, $p < 0.001$, *partial* $\eta^2 = 0.590$. This main effect was qualified by an interaction between learning condition and object similarity, $F(2,76) = 3.795$, $p = 0.027$, *partial* $\eta^2 = 0.091$ (see Fig. 10). Paired samples *t*-tests on the mean number of errors revealed similar patterns for each learning condition.¹ For participants in the haptic learning condition, objects that were 1 CB apart

(mean = 1.81 errors) were confused more often than objects that were 2 CBs apart (mean = 0.72 errors), $t(19) = 8.80$, $p < 0.001$, Cohen's $d = 1.801$, and 3 CBs apart (mean = 0.75 errors), $t(19) = 6.439$, $p < 0.001$, Cohen's $d = 1.76$. Objects that were 2 CBs and 3 CBs apart were confused equally often, $t(19) = -0.298$, *ns*. For participants in the visual learning condition, objects that were 1 CB apart (mean = 1.45 errors) were confused more often than objects that were 2 CBs (mean = 0.85 errors), $t(19) = 4.461$, $p < 0.001$, Cohen's $d = 0.820$, and 3 CBs apart (mean = 0.788), $t(19) = 4.715$, $p < 0.001$; Cohen's $d = 0.909$. Objects that were 2 CBs and 3 CBs apart were confused equally often, $t(19) = 0.531$, *ns*. Essentially, the larger effect sizes suggest that the interaction arose from a larger effect of object similarity on haptic identification than on visual identification.

Experimental phase

Learning condition, testing modality and congruence

Reaction times to correct trials were trimmed recursively at 3 standard deviations (this resulted in the removal of less than 0.5% of the data), averaged, and entered in a 2 (learning condition) \times 2 (testing modality) \times 2 (congruence) ANOVA, where the between factor was learning condition, and the repeated factors were testing modality and congruence. There was a main effect of learning condition, $F(1, 38) = 5.471$, $p = 0.025$, *partial* $\eta^2 = 0.126$. Generally, participants in the haptic learning condition (mean RT = 3,159 ms) responded faster than participants in the visual learning condition (mean RT = 3,856 ms). There was also a main effect of testing modality $F(1,38) = 19.59$, $p < 0.001$, *partial* $\eta^2 = 0.340$. These main effects were qualified by two interactions. There was an interaction between learning condition and testing modality, $F(1,38) = 5.904$, $p = 0.020$, *partial* $\eta^2 = 0.134$. Repeated-measures *t*-tests confirmed that, as was observed in Experiment 1, participants in the visual learning condition were faster at identifying objects visually (mean RT = 3,596 ms) than at identifying objects haptically (mean RT = 4,117 ms) $t(19) = 4.709$, $p < 0.001$, Cohen's $d = 0.461$. For participants in the haptic learning condition, there was no difference between visual identification (mean RT = 3,080 ms) and haptic identification (mean RT = 3,235 ms), $t(19) = 1.456$, *ns*. Finally, there was an interaction between testing modality and congruence, $F(1,38) = 8.304$, $p = 0.006$, *partial* $\eta^2 = 0.179$ (see Fig. 11). Bonferroni-protected *t*-tests revealed that when participants identified objects haptically, congruent trials (mean RT = 3,745 ms) were slower than incongruent trials (mean RT = 3,606 ms), $t(19) = 2.654$, $p = 0.011$, Cohen's $d = 0.26$. In contrast, when participants identified objects visually, congruent trials (mean RT = 3,299 ms) were not significantly slower than incongruent trials (mean RT = 3,381 ms), $t(19) = -1.387$, *ns*.

¹ To be consistent with the analyses of the three-way interaction observed in Experiment 1, we first compared visual and haptic identification at each level of object similarity. However, none of the comparisons were significant.

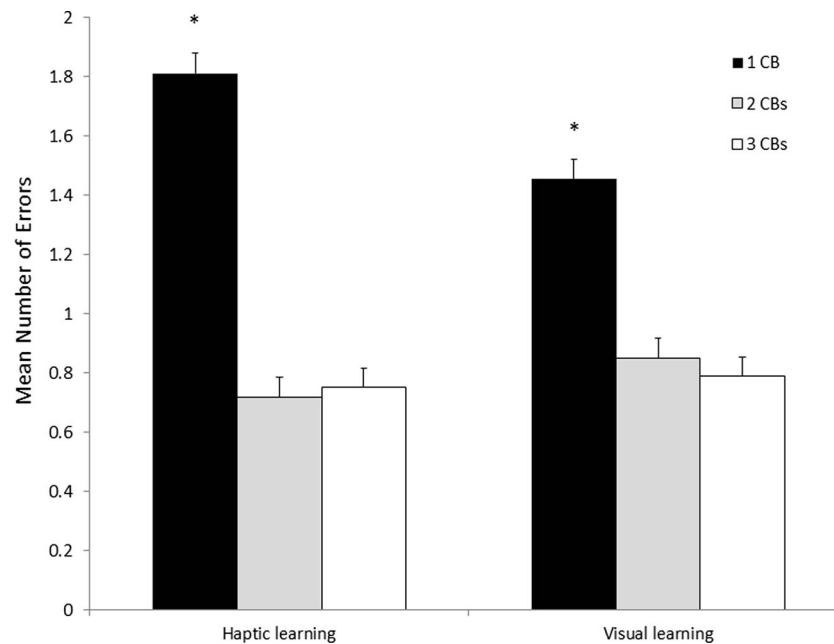


Fig. 10 Mean number of errors and standard error between objects that are 1 CB, 2 CBs, and 3 CBs apart for participants in the haptic and visual learning conditions. Asterisks denote statistically significant differences between objects that are 1 CB apart vs objects that are 2 or 3 CBs apart

We then entered the arsin of the square root of the mean proportion of errors in a 2 (learning condition) \times 2 (testing modality) \times 2 (congruence) ANOVA, where the between factor was learning condition, and the repeated factors were testing modality and congruence. We observed a main effect of congruence, $F(1,38) = 21.551$, $p < 0.001$, $partial \eta^2 = 0.362$. Participants were more accurate during congruent trials

(mean = 0.091) than during incongruent trials (mean = 0.120). We also observed a main effect of testing modality, $F(1,38) = 28.434$, $p < 0.001$, $partial \eta^2 = 0.428$, which was qualified by an interaction between learning condition and testing modality, $F(1,38) = 11.101$, $p = 0.002$, $partial \eta^2 = 0.226$. Directional repeated measures t -tests confirmed that participants in the visual learning condition produced more

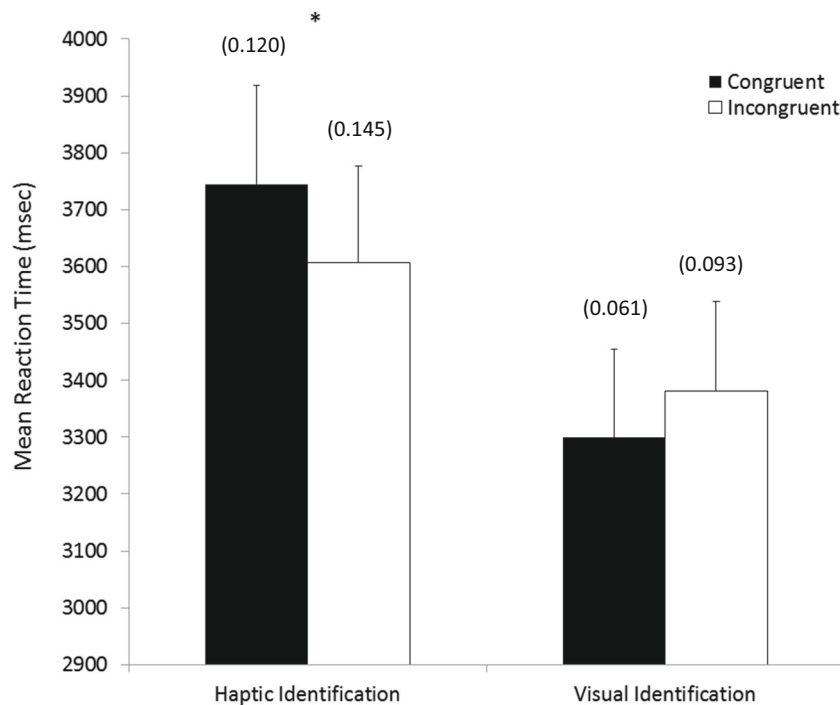


Fig. 11 Mean reaction times (arsin of square root of mean proportion of errors) and standard error for the haptic and visual identification of objects during congruent and incongruent trials. The asterisk denotes significant differences in reaction times between congruent and incongruent trials

errors when identifying objects haptically (mean = 0.148) than when identifying objects visually (mean = 0.058), $t(19) = 4.811$, $p < 0.001$, Cohen's $d = 1.302$. Again, participants in the haptic learning condition did not produce more errors when identifying objects visually (mean = 0.097) than when identifying objects haptically (mean = 0.118), $t(19) = 2.30$, *ns*.

Object similarity

Reaction times to correct incongruent trials were entered a 2 (learning condition) \times 2 (testing modality) \times 3 (objects similarity) mixed design ANOVA, where the between subject factor was learning condition, and the repeated factors were testing modality and object similarity. The results are presented in Figs. 12 and 13. There was a main effect of learning condition, $F(1,38) = 4.210$, $p = 0.047$, *partial* $\eta^2 = 0.100$. Participants in the haptic learning condition (mean = 3,187 ms) identified objects faster than participants in the visual learning condition (mean = 3,789 ms). This main effect was qualified by two interactions. First, there was an interaction between learning condition and testing modality, $F(1,38) = 7.765$, $p = 0.008$, *partial* $\eta^2 = 0.170$. Planned paired-sample *t*-tests confirmed that the pattern of results paralleled those observed in Experiment 1. For participants in the haptic learning condition, there was no difference between haptic identification (mean = 3,135 ms) and visual identification (mean = 3,149 ms) $t(19) = -0.100$, *ns*. For participants in the visual learning condition, haptic identification (mean = 4,078 ms) was slower than visual identification (mean = 3,613 ms), $t(19) = 4.352$, $p < 0.001$, Cohen's $d = 0.426$. There was also an interaction between learning condition and object similarity, $F(2,76) = 7.737$, $p = 0.008$, *partial* $\eta^2 = 0.169$. For participants in the visual learning condition, identifying objects when the distracter was 1 CB apart (mean = 4,017 ms) was slower than when the distracter was 2 CBs apart (mean = 3,758 ms), $t(19) = 3.393$, $p < 0.005$, Cohen's $d = 0.209$ or 3 CBs apart (mean = 3,593 ms), $t(19) = 3.964$, $p < 0.001$, Cohen's $d = 0.242$. The identification of objects when the distracter was 2 CBs apart was also slower than when the distracter was 3 CBs apart, $t(19) = 2.070$, $p < 0.05$, Cohen's $d = 0.163$. In contrast, for participants in the haptic learning condition, the identification of objects when the distracter was 1 CB apart (mean = 3,071 ms), 2 CBs apart (mean = 3,146 ms) and 3 CBs apart (mean = 3,345 ms) did not differ: $t(19) = -1.064$, *ns* for 1 CB versus 2 CBs, $t(19) = -1.42$, *ns* for 1 CB versus 3 CBs, and $t(19) = -1.171$, *ns* for 2 CBs versus 3 CBs. Object similarity impacted reaction time for participants who learned about the objects visually, but not for participants who learned about the objects haptically.

Finally, the arcsin of the square root of the mean proportion of errors during incongruent trials were entered in a 2 (learning condition) \times 2 (testing modality) \times 3 (object similarity) mixed

design ANOVA, for which the between factor was learning condition and the repeated factors were testing modality and object similarity. There was a main effect of testing modality, $F(1,38) = 14.730$, $p < 0.001$, *partial* $\eta^2 = 0.279$, and a main effect of object similarity, $F(2,76) = 17.842$, $p < 0.001$, *partial* $\eta^2 = 0.320$. No other effects were significant, but there was a trend towards an interaction between learning condition and testing modality, $F(1,38) = 2.986$, $p = 0.092$, *partial* $\eta^2 = 0.073$. Planned pairwise comparisons confirmed that for participants in the haptic learning condition there was no difference between haptic identification (mean = 0.132) and visual identification (mean = 0.107), $t(19) = 1.574$, *ns*. As expected, participants in the visual learning condition made more errors during haptic identification (mean = 0.159) than during visual identification (mean = 0.079), $t(19) = 3.855$, $p = 0.001$, Cohen's $d = 1.086$.

General discussion

We contrasted haptic and visual processing by asking participants to learn to recognize novel objects haptically, visually, or bimodally, and then asked them to identify eight novel objects visually and haptically. We then presented them simultaneously with two objects (one haptically presented, and one visually presented) and asked them either to identify one of the two objects or to indicate whether the two objects were the same. Our results demonstrate three important points. First, we showed that haptic identification and visual identification follow similar principles. Second, we provided evidence that performance is not always best when learning and testing conditions match. Though this was true for participants who learned to recognize objects visually or bimodally, participants who learned to recognize objects haptically performed equally well when tested haptically or visually. Finally, we showed that when bimodal information is presented for matching or identification, object similarity can impact performance. Object similarity impacted multisensory performance when matching objects across modalities, and when identifying objects—but only when participants initially learned to recognize objects visually.

Shared memory representations for haptic and visual inputs

Using a set of psychophysically scaled objects, we demonstrated that haptic memory representations and visual memory representations are quite similar. In both modalities, similar objects were confused more often than distinct objects, and confusions occurred more often on the curvature dimension. Importantly, the patterns of errors across identification modalities were correlated; individual participants confused the

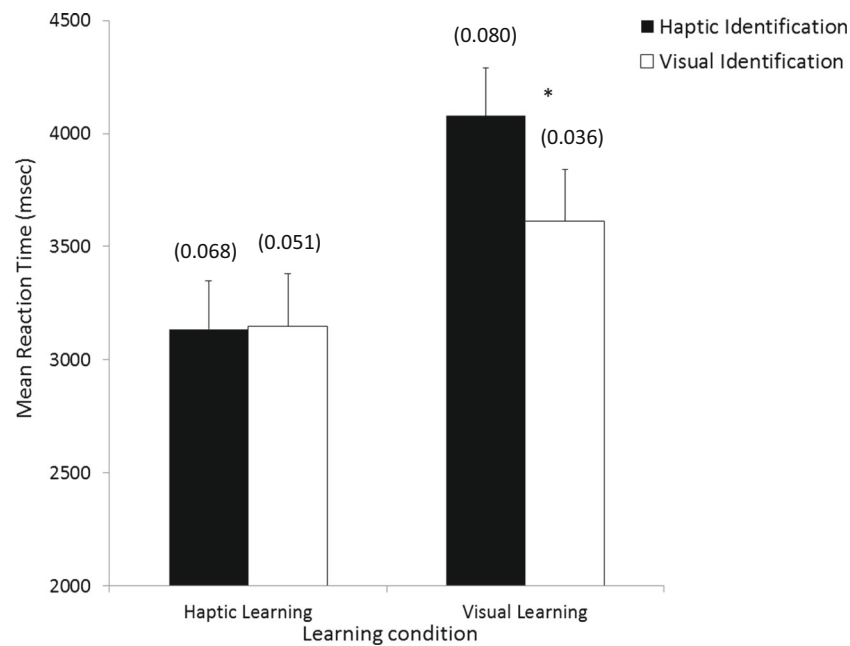


Fig. 12 Mean reaction time (arsin of square root of mean proportion of errors) and standard error for the haptic identification and visual identification of participants in the haptic and visual learning conditions. Asterisks indicate a statistically significant difference between haptic and visual identification

same two objects when they attempted haptic or visual identification (see Loomis 1981, 1982 for similar results). Though this finding is not surprising, it establishes that haptic and visual identification are directly comparable using this stimulus set.

The observation that the memory representations developed from haptic and visual inputs are similar is consistent with two potential explanations. One, haptic and visual representations, though distinct representations, may rely on very

similar or identical principles—at least in terms of how they represent shape information. A second, more parsimonious explanation would be that haptic and visual identification may rely on a shared representation that was created from both initial inputs. Indeed, the notion that both inputs accessed a single memory representation explains that the patterns of results are parallel for the two input modalities. Results were highly similar for haptic and visual identification because both tasks relied on the same memory representations. Our studies

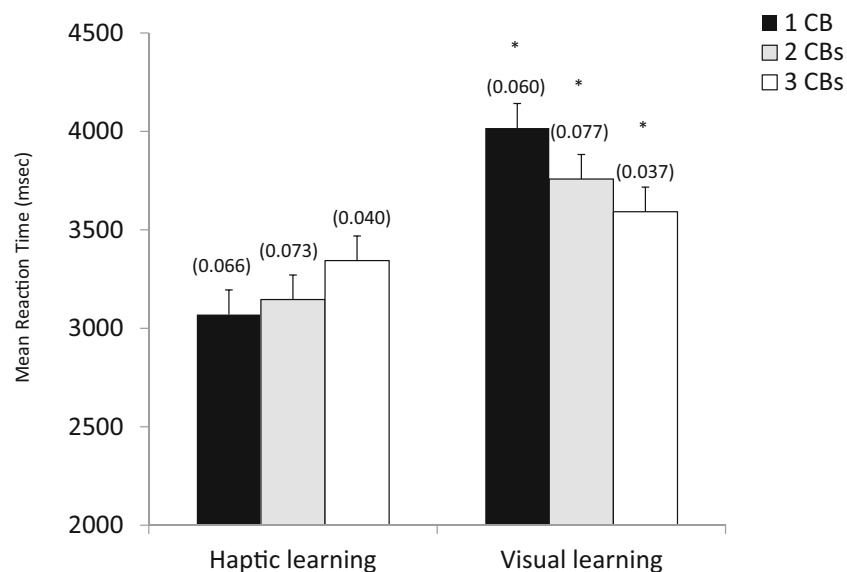


Fig. 13 Mean reaction time (arsin of square root of mean proportion of errors) and standard errors for the identification of objects when distracters are 1 CB, 2 CBs, and 3 CBs apart for participants in the

haptic learning condition and the visual learning condition. Asterisks indicate reaction times that are significantly different from each other

did not directly evaluate this possibility, but the interference patterns observed are consistent with the idea that both tasks were based on a shared representation.

In Experiment 1, performance was generally slower during incongruent trials than during congruent trials. This occurred for haptic identification, where (consistent with Pavani et al., 2000, Rock and Viktor, 1964, and Spence et al., 2004) incongruent visual information interfered with haptic processing. Importantly, this occurred despite the fact that participants began reaching for an object 1 s before opening their eyes. Visual information presented shortly *after* haptic information still slowed down responses. We also go beyond the findings of Pavani et al., (2000) and Spence et al., (2004) and showed that this also occurred for visual identification as well. Incongruent haptic information interfered with visual identification. When an irrelevant object was presented (the object that was not to be identified), its presentation seemed to automatically activate its representation in memory, and this representation interfered with the identification of the target object. If both the haptic and the visual inputs activated the memory representations in the same network, these representations would interfere with each other and slow reaction times. A common memory representation is therefore consistent with interference, as well as the fact that we saw interference despite participants' knowledge that there were two separate objects, and despite instructions to ignore the irrelevant object.

If vision and touch share memory representations, why did we not see longer reaction times for incongruent trials in Experiment 3? In Experiment 3, the effect of congruence was limited. Participants were slower during congruent trials compared to incongruent trials, but only during haptic identification. Furthermore, this effect on reaction time was accompanied by a main effect on error rates where participants were generally more accurate during congruent trials, suggesting a potential speed-accuracy trade-off. Therefore, it would be rash to state that performance was better during incongruent trials, and we will instead discuss this finding from the point of view that there was no clear interference during incongruent trials. Reaction times were generally slower in Experiment 3 compared to Experiment 1. This would be consistent with our earlier suggestion that, in Experiment 1, asking participants to complete both haptic identification trials and visual identification trials led to richer representations in memory, which led to faster reaction times and a lack of impact from object similarity. In Experiment 3, participants initially identified objects only haptically or only visually—potentially leading to less elaborate representations in memory (at least for participants in the visual learning condition), and to overall slower reaction times. Furthermore, the lack of familiarity with the unpracticed modality may have further slowed reactions times. Consequently, the slower initial processing may have

left little space for an overall slowing of reaction times due to incongruent distracters. However, it left enough space for this slower reaction time to be modulated by the similarity of the representations in memory between the object to be named and the distracter—at least for participants who learned about objects visually who may have had less elaborate representations in memory.

The effect of object similarity that we observed in Experiment 2 (for the matching task) and Experiment 3 (for participants who learned to recognize objects visually) further supports the idea that the object representations were shared. According to object identification models such as the Hierarchical Interactive Theory (HIT) model (Humphreys & Forde, 2001), when an object is presented for identification, it activates its representation in memory as well as the representations of other similar objects, in a spreading activation fashion. Identification occurs when the representation of an item reaches a specific threshold. In our paradigm, two objects were presented at once. When both objects are the same (during congruent trials), the same representations in memory are activated, and threshold is reached relatively quickly. However, when the two presented objects are different, two separate representations are activated, one for each object. When two distinct objects were presented (for example, two objects that were 3 CBs apart), there is not much competition for activation, and threshold can be reached relatively quickly, leading to faster reaction times. However, when two similar objects are presented (for example, two objects that were 1 CB apart), the two representations compete for activation, leading to longer reaction times. Note that both inputs activate representations in the same memory system. Altogether, our results strongly support the idea of shared representations.

Within modality advantage?

Contrary to previous findings, our results demonstrate that performance is not always best when learning conditions and testing conditions match. One pattern that was consistent across experiments was that participants who learned to recognize objects visually (whether strictly visually or within a bimodal learning condition) performed better when identifying objects visually than when identifying objects haptically (*see* Easton, Greene, and Srinivas, 1997; Newell et al., 2005; as well as Norman et al., 2004 for similar results). Interestingly, when participants learned to recognize objects haptically, performance was equivalent across haptic and visual identification. Participants who learned to recognize objects by touch were just as accurate when asked to recognize the same objects by sight (*see* Easton, Srinivas and Green, 1997 for similar findings).

Why would learning about objects haptically lead to equivalent performance in haptic and visual identification? It is

possible that participants who explore stimuli haptically generate a visual memory trace as well as a haptic memory trace. When participants learn to recognize objects haptically, they can visualize what the objects would look like, leading to a visual memory trace that would resemble the one created when objects are seen. In fact, Lacey and Campbell (2006) observed a disruption of their participants' ability to recognize novel objects that had been haptically explored when asked to look at a visual display that showed irrelevant noise, suggesting that a visual representation was created from the haptic input. The fact that a visual memory trace is created from the haptic input would explain the equivalent performance in haptic and visual identification for individuals who first learned to recognize the objects haptically. Individuals who first learned to recognize the objects visually may not create this dual trace: it is arguably easier to imagine what a held object would look like than it is to imagine what a seen object would feel like in our hand. Here we are not suggesting two separate memory representations, but rather a single memory representation that comprises multiple and sometimes overlapping pieces of information.

A second possibility is that participants who first learned to recognize objects haptically used verbal labels to describe each object, creating more elaborate memory representations that are easily accessed from a haptic or a visual input. Past studies that reported better within-modality performance used scenes (Newell et al., 2005) or complex pepper-shaped objects (Norman et al. 2004), stimuli that may not easily be described verbally. In contrast, our stimulus set comprised objects that were much simpler—they varied on three predetermined dimensions—and that could easily be verbally described using these dimensions. The ease with which we can verbally describe stimuli may therefore influence the kind of representation that is constructed in memory. There are indeed other findings that confirm that items that are verbalized more easily can also be remembered or recognized more easily (Desmarais, Lane, LeBlanc, Hiltz, & Richards, 2017; Silverberg & Buchanan, 2005), and that a verbal mask can interfere with object processing (Desmarais, Li, & Anderson, 2017; Lacey & Campbell, 2006). Haptically explored stimuli that could be verbally described, such as our stimulus set, may therefore generate more elaborate memory representations that can be more easily accessed for visual identification. We may not see such facilitation for participants who were first asked to learn to recognize objects visually because visual perception may seem inherently easier—the form of the object is readily accessible. Though this is speculative, it is possible that the verbalization described above is done during haptic exploration in order to enhance performance. It is also possible that the limitation to 2 s given to participants for visual exploration

does not give them enough time to verbalize the form of the objects, but recall that participants in the bimodal condition were given a full 5 s to explore objects. If the time given for exploration determined whether verbalization was possible, participants in the bimodal condition should have been able to verbalize. The fact that their pattern of performance mirrored that of participants in the visual learning condition suggests that time was not the determining factor.

Are multimodal inputs better?

When we asked participants to identify multisensory stimuli, we observed that participants in the bimodal learning condition tended to perform more poorly than participants in unimodal learning conditions when asked to identify objects haptically. These participants were exposed to haptically presented objects during learning trials, yet they identified these objects more slowly and less accurately. This is consistent with past research that shows worse performance during bimodal trials (i.e., responding to unimodal trials vs bimodal trials; see Colavita, 1974; Hecht & Reiner, 2009; Sinnott, Soto-Faraco, & Spence, 2008). In the traditional Colavita task, participants identify the modality that target stimuli are presented in: auditory, visual, or bimodal. When bimodal stimuli are presented, participants tend to *fail to detect* the auditory stimulus and report only the presence of the visual stimulus, either with neutral stimuli like lights and tones (Colavita, 1974), with a thin gray bar and a mechanical resisting force (Hecht & Reiner, 2009), or with line drawings and sounds of common objects (Sinnott et al., 2008). The Colavita task suggests that vision is dominant, and that it can lead participants to fail to detect an auditorily presented stimulus or a haptically presented stimulus. In these studies, it is the currently presented stimulus that is not detected. In our study, it was prior exposure during the learning phase that led to poorer performance when identifying haptically presented objects, when the visually presented distracter was incongruent. However, a similar process may have been at work.

Recall that during learning trials in the bimodal learning condition, one object was presented haptically, and another identical object was presented visually. Because there was a visually-presented object, there was no need to focus on the haptically presented object or to visualize it. This would be akin to the Colavita effect where participants fail to detect the non-dominant stimulus. The overall similarity in performance between the participants in the visual and bimodal learning conditions supports the idea that preference was given to the visual input in bimodal learning trials. The result of this inattention to the haptic stimulation during learning trials can then lead to a weaker or noisier haptic memory trace—despite the fact that participants were told that they would later have to

identify objects haptically. This weaker haptic memory trace may not be problematic during the unimodal haptic trials of Experiment 1 because visual and haptic test trials were blocked, and the haptically presented objects could still be visualized because there was no competing visual input. Performance was consequently similar to that of participants in the unimodal visual learning condition. This same noisier memory trace could, however, create the accuracy differences observed in the bimodal stimuli of the experimental phase. Here, participants were presented with two objects: one haptically and one visually, creating two competing inputs. Furthermore, haptic and visual trials were mixed in order to prevent participants from focusing on a single modality and ignore the other. It is possible that participants in the bimodal learning condition, having suppressed the haptic input during learning trials, focused more heavily on the visual input. This could then result in reaction times similar to those observed for participants in the visual learning condition, but in a lower accuracy—as was observed here. Altogether, the pattern of results suggests that the presence of visual information during bimodal learning trials may serve to suppress the encoding of haptic information. This would explain why the advantage observed for participants in the haptic learning condition (equivalent performance for haptic and visual identification) was not observed for participants in the bimodal learning condition.

Summary and conclusion

We compared haptic and visual identification by asking participants to learn to recognize novel objects that varied only in shape. We demonstrated that both haptic and visual identification relied on similar processes, that patterns of errors are very similar across modalities, and that incongruent information, even if irrelevant to the task, can impact object identification. Our results are consistent with the idea that haptic and visual identification draw on shared representations.

Acknowledgements The authors would like to thank Claire Whittaker and Joey Anderson for their assistance. This research was supported by a Marjorie-Young Bell faculty fellowship awarded to the first author.

References

- Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, *391*, 756.
- Bryant, P. E., & Raz, I. (1975). Visual and tactual perception of shape by young children. *Developmental Psychology*, *11*(4), 525–526.
- Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409–412.
- Desmarais, G., & Dixon, M. J. (2005). Understanding the structural determinants of object confusion in memory: An assessment of psychophysical approaches to estimating visual similarity. *Perception & Psychophysics*, *67*(6), 980–996.
- Desmarais, G., Dixon, M., & Roy, E. A. (2007). A role for action knowledge in visual object identification. *Memory & Cognition*, *35*(7), 1712–1723.
- Desmarais, G., Lane, B., LeBlanc, K. A., Hilts, J., & Richards, E. D. (2017). What’s in a name? The influence of verbal labels on action production in novel object/action associations. *Visual Cognition*. doi:10.1080/13506285.2017.1308451
- Desmarais, G., Li, Q., & Anderson, W. (2017). What interference reveals about object representations in haptic and visual identification. Poster presented at the 27th annual meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, University of Regina, Regina, 2–4 June 2017.
- Dickerson, J., & Humphreys, G. W. (1999). On the identification of misoriented objects: Effects of task and level of stimulus description. *European Journal of Cognitive Psychology*, *11*, 145–166.
- Easton, R. D., Greene, A. J., & Srinivas, K. (1997). Transfer between vision and haptics: Memory for 2-D patterns and 3-D objects. *Psychonomic Bulletin & Review*, *4*(3), 403–410.
- Easton, R. D., Srinivas, K., & Greene, A. J. (1997). Do vision and haptics share common representations? Implicit and explicit memory within and between modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(1), 153–163.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.
- Ernst, M. O., & Bühlhoff, H. H. (2004). Merging the sense into a robust percept. *Trends in Cognitive Neurosciences*, *8*(4), 162–169. doi:10.1016/j.tics.2004.02.002
- Gaissert, N., & Wallraven, C. (2012). Categorizing natural objects: A comparison of the visual and haptic modalities. *Experimental Brain Research*, *216*(1), 123–134.
- Gallace, A., & Spence, C. (2009). The cognitive and neural correlates of tactile memory. *Psychological Bulletin*, *135*(3), 380–406. doi:10.1037/a0015325
- Hecht, D., & Reiner, M. (2009). Sensory dominance in combinations of audio, visual, and haptic stimuli. *Experimental Brain Research*, *193*, 307–314. doi:10.1007/s00221-008-1626-z
- Heller, M. A. (1983). Haptic dominance in form perception with blurred vision. *Perception*, *12*, 607–613.
- Hershberger, W. A., & Misceo, G. F. (1996). Touch dominates haptic estimates of discordant visual-haptic size. *Perception & Psychophysics*, *58*(7), 1124–1132.
- Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: “Category-specific” neuropsychological deficits. *The Behavioral and Brain Sciences*, *24*, 453–509.
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, *5*, 67–103.
- Klatzky, R. L., Lederman, S. J., & Metzger, V. A. (1985). Identifying objects by touch: An expert system. *Perception & Psychophysics*, *37*(4), 299–302.
- Lacey, S., & Campbell, C. (2006). Mental representation in visual/haptic crossmodal memory: Evidence from interference effects. *The Quarterly Journal of Experimental Psychology*, *59*(2), 361–376. doi:10.1080/17470210500173232
- Lacey, S., Pappas, M., Kreps, A., Lee, K., & Sathian, K. (2009). Perceptual learning of view independence in visuo-haptic object representations. *Experimental Brain Research*, *198*(2–3), 329–337. doi:10.1007/s00221-009-1856-8
- Lloyd-Jones, T. B., & Humphreys, G. W. (1997a). Perceptual differentiation as a source of category effects in object processing: Evidence from naming and object decision. *Memory & Cognition*, *25*, 18–35.

- Lloyd-Jones, T. B., & Humphreys, G. W. (1997b). Categorizing chairs and naming pears: Category differences in object processing as a function of task and priming. *Memory & Cognition*, *25*, 606–624.
- Loomis, J. M. (1981). On the tangibility of letters and braille. *Perception & Psychophysics*, *29*(1), 34–46.
- Loomis, J. M. (1982). Analysis of tactile and visual confusion matrices. *Perception & Psychophysics*, *31*(1), 41–52.
- Mancini, F., Bricolo, E., & Vallar, G. (2010). Multisensory integration in the Müller-Lyer illusion: From vision to haptics. *The Quarterly Journal of Experimental Psychology*, *63*(4), 818–830.
- Newell, F. N., Woods, A. T., Mernagh, M., & Bühlhoff, H. H. (2005). Visual, haptic, and crossmodal recognition of scenes. *Experimental Brain Research*, *161*, 233–242. doi:10.1007/s00221-004-2067-y
- Norman, J. F., Norman, H. F., Clayton, A. M., Lianekhammy, J., & Sielke, G. (2004). The visual and haptic perception of natural object shape. *Perception & Psychophysics*, *66*(2), 342–351.
- Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological Science*, *11*(5), 353–359.
- Peltier, S., Stilla, R., Mariola, E., LaConte, S., Hu, X., & Sathian, K. (2007). Activity and effective connectivity of parietal and occipital cortical regions during haptic shape perception. *Neuropsychologia*, *45*, 476–483. doi:10.1016/j.neuropsychologia.2006.03.003
- Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, *143*, 594–596.
- Shepard, R. N. (1987). Towards a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.
- Silverberg, N., & Buchanan, L. (2005). Verbal mediation and memory for novel figural designs: A dual interference study. *Brain and Cognition*, *57*, 198–209. doi:10.1016/j.bandc.2004.08.045
- Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, *128*, 153–161. doi:10.1016/j.actpsy.2007.12.002
- Spence, C., Pavani, F., & Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 148–169.
- Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology: A*, *47*(3), 631–650. doi:10.1080/14640749408401131