Isolating shape from semantics in haptic-visual priming

Ana Pesquita[1], Allison A. Brennan[1], James T. Enns[1], Salvador Soto-Faraco[2]

[1]University of British Columbia, Department of Psychology

2136 West Mall V6T 1Z4 Vancouver BC, Canada

[2]Universitat Pompeu Fabra

Departament de Tecnologies de la Informació i les Comunicacions

Carrer Tànger 122-14008018 Barcelona, Spain

Address correspondence to:

anapesquita@psych.ubc.ca

Tel: (+1)7789871982

Fax: (+1)6048226923

Abstract

The exploration of a familiar object by hand can benefit its identification by eye. What is unclear is how much this multisensory cross-talk reflects shared shape representations versus generic semantic associations. Here we compare several simultaneous priming conditions to isolate the potential contributions of shape and semantics in haptic-to-visual priming. Participants explored a familiar object manually (haptic prime) while trying to name a visual object that was gradually revealed in increments of spatial resolution. Shape priming was isolated in a comparison of identity priming (shared semantic category and shape) with category priming (same category, but different shapes). Semantic priming was indexed by comparisons of category priming with unrelated haptic primes. The results showed that both factors mediated priming, but that their relative weights depended on the reliability of the visual information. Semantic priming dominated in Experiment 1, when participants were free to use high-resolution visual information, but shape priming played a stronger role in Experiment 2, when participants were forced to respond with less reliable visual information. These results support the structural description hypothesis of haptic-visual priming (Reales and Ballesteros, 1999) and are also consistent with the optimal integration theory (Ernst and Banks 2002), which proposes a close coupling between the reliability of sensory signals and their weight in decision making.

Keywords: multisensory, vision, haptic, crossmodal, priming, weighted decisions

When you reach for a flashlight in your backpack at dusk, both its seen and felt properties help you distinguish it from your camping knife. In this case, information that is uncertain in either vision or touch considered alone tends to be combined to help us form a more robust representation (Ernst and Bülthoff 2004; Newell 2004; Welch and Warren 1986). Although previous research has documented these cross-modal benefits for haptics and vision, the nature of this inter-sensory influence is less clear. This is because, as in the foregoing example, reaching for the flashlight not only triggers shape perception, but it activates rich semantic associations that can help identify the flashlight (Neely 1977; Kahlaoui et al. 2007). Here we isolate the contribution of shared shape information from the contribution of semantic associations in the haptic-to-visual priming of familiar objects.

Previous research indicates that the identification of familiar visual objects (e.g., pipe, lock) is facilitated by congruent haptic information (Reales and Ballesteros 1999). Reales and Ballesteros interpreted this haptic-to-vision priming as implicating the cross-modal transfer of shape information between touch and vision. However, this study has been criticized because it used familiar objects (Craddock and Lawson 2008; Ernst et al. 2007; James et al. 2002). Testing familiar objects raises the possibility that some or all of the cross-modal benefit was conceptually mediated, since familiar objects activate long-term memory representations via both haptic and visual modalities (Humphreys et al. 1999; Klatzky and Lederman 1999). In an effort to rule out semantic priming, Reales and Ballesteros reported priming benefits when semantic labeling was not an explicit requirement of the task (object versus non-object discrimination). However, this manipulation does not prevent semantic priming conclusively if object labeling can occur involuntarily (Glaser 1992, Piai et al. 2011, Tipper and Driver 1988).

Another approach to isolate shape from semantics in cross-modal priming involves manipulating the size (Craddock and Lawson 2009) and orientation in depth (Lawson 2009) of familiar object primes and targets. These studies show that priming is sensitive to these object properties, independent of any semantic influences. However, these studies did not assess how shape and semantic information might *each* contribute to the priming of familiar object recognition and how their relative contributions might vary as a function of the signal strength in each modality.

Other studies have used unfamiliar three-dimensional shapes in an attempt to rule out semantic priming as an explanation (Ernst and Banks 2002; Ernst et al. 2007; Grefkes et al. 2002; James et al. 2002; Newell et al. 2001; Rentschler et al. 2004). However, even unfamiliar shapes may be rapidly assigned to verbal-semantic categories during the course of experiment. For example, abstract clay objects may be categorized into those that look more like 'boats' or 'plants' (James et al. 2002). Such categorization would activate semantic associations, even though the stimuli were novel. In support of this possibility, Lacey and Campbell (2006) reported that a secondary verbal task paired with a primary haptic-visual priming task interfered more with priming of unfamiliar than of familiar objects. This implies that participants strategically encode properties of unfamiliar shapes using verbal descriptors. Moreover, these studies do not address the question of how shape and semantics may both be involved in the cross-modal identification of *familiar* objects.

Evidence for shape and semantic contributions to priming

Before describing our approach to isolating shape from semantics, it is important to point to the growing evidence for both types of influence in haptic-visual priming. Some of the evidence of shared shape representations comes from brain imaging studies. For instance, the lateral occipital complex (LOC) is activated by objects seen either visually or grasped by hand (Amedi et al. 2001; Amedi et al. 2005; Lacey et al. 2009). Other brain regions that have been implicated in cross-modal shape transfer include the insula (Grefkes et al. 2002; Hillis and Caramazza 1995; Saito et al. 2003), and the intraparietal sulcus (IPS; Grefkes et al. 2002; James et al. 2002). Several studies have shown that the insula is involved in registering simple features (Grefkes et al. 2002; Hillis and Caramazza 1995; Saito et al. 2003), while the IPS and the LOC seem more involved in the representation of volumetric shape (Grefkes et al. 2002; James et al. 2002). In addition, the LOC plays a role in the association of the multisensory percept with semantic knowledge, as occurs during familiar object recognition (Amedi et al., 2001).

There is also evidence for multisensory shape representations in developmental studies. For example, pre-verbal infants look longer at shapes they have previously experienced through oral contact (Gottfried et al. 1977; Meltzoff and Borton 1979; Rose 1994). Although verbal mediation is effectively ruled out as a priming mechanism in infants, it is still possible that verbal-conceptual knowledge is unavoidably involved in visual-haptic priming for older participants. Research with the elderly also shows robust haptic-visual priming (Ballesteros et al. 2009), though separating shape and semantic contributions was not the focus of that work.

Automatic semantic associations are widely reported and appear to be activated in response to any input modality (Carr et al. 1982, Easton, Srinivas, & Greene, 1997; Easton, Greene, & Srinivas, 1997; Friese et al. 2012, Orgs et al. 2006, Schacter et al. 2004, Stadthagen-Gonzalez et al. 2009; Tipper and Driver, 1988). For example, the identification of familiar objects in pictures is influenced by written words that cause the activation of related semantic concepts (Kahlaoui et al. 2007; Neely 1977). Several studies indicate that objects that are named more easily tend to show larger cross-modal effects. For example, Bushnell and Baxt (1999) reported that children's visual recognition of objects that had been previously explored haptically was poorer for unfamiliar than familiar objects. Lacey and Campbell (2006) reported that familiar objects were recognized more accurately than unfamiliar ones in a visual-haptic memory task. Craddock (2008) observed that performance was both faster and more accurate for familiar than unfamiliar objects in a visual-haptic recognition task. Taken together, these findings suggest that the activation of semantic knowledge in one modality has an influence on perceptual decisions made in another modality.

The possibility that semantic associations assist visual-haptic priming in human adults has been formalized in the dual-coding theory of Johnson et al. (1989; see also Paivio 1991). According to this theory, a familiar haptic input activates a visually coded shape representation of the felt object, which in turn activates a corresponding semantic representation. In Johnson et al. (1989) children were tested for visuospatial and verbal abilities before performing a cross-modal shape recognition task. Task instruction was a critical factor. The performance of children who had been asked to complete the cross-modal recognition task using imagery correlated with their visuospatial ability; in contrast, the performance of children who followed naming instructions correlated with verbal skill. When given no instructions on how to complete the task, children tended to use a naming strategy. The implication is that object perception automatically and involuntarily triggers semantic associations (Glaser 1992; Piai et al. 2011), making it difficult to single out the possible contribution of shape representations that are shared across sensory modalities.

Scope of the present study

The approach we used to separate the influences of shared shape representations from semantic associations was inspired in part by a visual-priming study of Biederman and Cooper (1991). These authors asked participants to name briefly presented and masked line drawings of familiar objects (e.g., a grand piano, to be called "piano"). Subsequently, participants then had to name either an identical test picture (i.e., grand piano, called "piano") or one showing a different shaped exemplar of the same basic category (e.g., upright-piano, called "piano"). The results showed that previously naming a prime picture facilitated the speed and accuracy of naming the test pictures, and that priming was greater for identical pictures than differently shaped exemplars. The authors interpreted the larger priming effect for identical over different exemplar images as evidence of visual shape priming, by virtue of the fact that any semantic activation of the object's label was a constant in these two conditions.[1]

---

[1] Biederman and Cooper (1991) emphasize that this comparison is, if anything, an underestimation of the full extent of shape priming, because different shaped exemplars are likely be partially similar to one another in shape. As such, the comparison is a lower-bounds estimate of the visual priming component and therefore also an upper-bounds estimate of semantic priming.

A second inspiration for our design came from Iordanescu et al. (2008). Participants in these experiments searched for a visual target object among distractor objects (e.g., keys among cats and toys) while simultaneously listening to sounds that were either characteristic of the target (e.g., a jingle) or of one of the distractors (e.g., a meow). The characteristic sound was always task-relevant, because it was congruent with one of the visual objects, but critically, it was never predictive. That is, sounds were no more likely to correspond to targets than any of the distractor objects. The results showed that successful target search was faster in the presence of a target-congruent sound than a distractor-congruent sound.

We combined these design features in a study of haptic-to-visual priming. Participants explored prime objects with their hands while a visual object was slowly revealed through the addition of higher spatial frequencies. Gradually revealing visual information allowed the two sensory modalities to interact while there was still uncertainty in each modality. The two measures of performance included: (1) accuracy in naming the visual object, and (2) the image resolution at which this successful identification could be made.

In an *identity* condition, the haptic prime and the visual target shared the same shape and semantic category label. To isolate shape-based priming, following Biederman and Cooper (1991), the identity condition was compared to a *category* condition, where the haptic prime and visual target objects had different shapes but shared the same semantic category label. Semantic plus shape priming was measured by comparing the category condition with an *unrelated* condition, where the haptic prime was a potential target but on this occasion shared neither the same shape nor label as the visual target. This comparison followed Iordanescu et al. (2008), in that it compared two conditions in which the haptic prime was task relevant (i.e., the haptic prime was a potential target) but in which no cross-modal prediction was possible (i.e., the haptic prime was no more likely to be the visual target than one of the non-targets on a given trial). Finally, to provide a baseline between these extremes, we included a *neutral* condition, where the haptic prime shared neither shape nor label with the visual target object, and was never a potential visual target. This condition was included to test for the possibility that holding a task-relevant object (i.e., as in the identical, category, and unrelated conditions) generally facilitated or interfered with visual identification. If so, the neutral condition should be different in accuracy from all three of these conditions. Alternatively, if congruent haptic objects facilitated visual identification and incongruent objects interfered with it, then accuracy in the neutral condition should fall between the category and unrelated conditions.

Three distinctly different patterns of outcome are possible for the identity, category, and unrelated conditions. First, accuracy may be higher in the *identity* than in the *unrelated* condition, but similar in the *category* and *unrelated* conditions. This would imply that haptic-visual priming is mediated solely by shared shape representations, as implied in the interpretation of Reales & Ballesteros (1999). Second, accuracy may be higher in both the *identity* and the *category* conditions, and to a similar extent, than accuracy in the *unrelated* condition. This result would imply that semantic associations are the sole influence of priming, and thus would agree with the critics of Reales & Ballesteros (1999). Finally, accuracy may be higher in the *identity* than in the *category* condition, which in turn is higher than in the *unrelated* condition. This pattern implies that both shape representations and semantic associations play a role in haptic-visual priming, consistent with dual-coding theories (Johnson et al., 1989; Paivio 1991).

Experiment 1

As illustrated in Figure 1A, participants actively explored a familiar object with both hands (haptic prime) while viewing a sequence of images that gradually revealed the picture of a familiar object (visual target). They were asked to name the visual target as soon as they could identify it. The visual target began as a highly blurred gray image and higher spatial frequencies were added until the native resolution was reached. The entire sequence took 15 sec (Online Resource 1). Our rationale for gradually revealing visual targets was that it gave the two sensory modalities an opportunity to interact while there was still uncertainty in each modality. This is important if sensory modalities combine information in a statistically optimal fashion, meaning that signals from different modalities are weighted according to their reliability (Ernst & Banks, 2002). If the modalities are not given this opportunity, one might falsely conclude that vision was always dominant over haptics, when instead it merely reached its threshold for certainty more rapidly.

Methods

Participants. Sixteen participants (10 female), between 21 and 33 years of age, from Universitat Pompeu Fabra (Barcelona, Spain) volunteered for the experiment. All were right-handed and had normal or corrected-to-normal visual acuity, and gave written informed consent. The procedures were approved by the local ethics committee (CEIC Parc de Salut Mar).

Stimuli and Apparatus. Sixteen familiar objects composed the stimuli set. This number of objects was constrained by the 1-hour testing session available to participants (allowing for 64 trials in all) and by our estimate of the minimal number of observations (16 trials) needed to get a stable estimate of accuracy for each participant in each of the four conditions. The objects were selected so that there were eight easily nameable semantic categories: brush, cup, strainer, mobile-phone, light bulb, comb, rock, and, stapler. Each category was represented in the set by 2 differently shaped exemplars (e.g., two different "light bulbs," two different "brushes," etc.), as shown in Figure 1B, so that they differed considerably in shape but yet each rapidly evoked the same verbal label. Visual presentation was done through a sequence of images that depicted an initially blurry object (low spatial frequencies only) gradually turning into a high-resolution image via a sequence of half-second still frames. In the neutral priming condition, participant's held a ping-pong ball, which they were told they would never see, because it was not a visual target on any trial.

The photos of each object were taken from 4 different vantage points, each corresponding to a common-use view of the object that did not hide any critical parts. The photos were viewed at a distance of 40 cm, subtended about 12° of visual angle and had a resolution of 260 x 260 pixels by 256 levels of gray. Each picture was filtered (using ImageJ v1.44 software, http://rsb.info.nih.gov/ij/) using spatial frequency bands that progressively increased in steps of 2 cycles/pixel. This procedure resulted in sequences of 30 different levels of filtering for each photo. Each sequence of filtered photos for a given object was played in reverse to participants, such that they began as highly blurred gray images, and after 30 frames the sequence ended with the photo in native resolution (Online Resource 1).

The testing apparatus consisted of a 50 cm tall x 50 cm wide x 80 cm deep wooden box, open at the front and back, and set on a table between the participant and the experimenter. The opening facing the participant was covered with a black cloth, with two vertical slits, to allow participants to reach inside the box with both arms, but not to see inside. The opening facing the experimenter was used to place objects inside the box, so that the participant could explore the objects with both hands. A black panel above the box visually separated the participant from the experimenter. On the floor of the box, was a button that participants pressed with their right index finger when they were in resting position (between trials). Underneath the table was a foot pedal for the participant's right foot. On top of the box, facing the participant, a CRT monitor (1,024 X 768 pixels, 75 Hz refresh rate) was used to display the visual sequences (Online Resource 2). The finger button, foot pedal, and visual displays were controlled by a computer running Matlab 2010a software and Psychtoolbox 3 library, and the interfaces used LabJack U-12.

Procedure. Each participant began with a familiarization phase where all the objects were explored by touch, in a random order. After placing each object in the participant's hands, the experimenter named the object using its category label, and then permitted 10 s for haptic exploration. Instructions emphasized that exploration be active and involve both hands. The familiarization phase was intended to give participants fluency with all the objects, so that name finding did not add unnecessary variance to the measurements.

Each experimental trial began with participants depressing the table button and the foot pedal with their right index finger and toes, respectively. During this time the experimenter placed an object inside the box and triggered a video message telling participants to pick up the object and actively explore it with both hands. In doing so, participants released the finger button, which triggered the onset of a visual de-blurring sequence on the computer screen. The primary task of the participant was to indicate as soon as they recognized the visual object by lifting off their foot of the pedal. As soon as a foot response was detected the visual image was erased from the screen. It was also the measure of object identification time used in our primary data analysis. The participant was then asked to name the visual object (the experimenter recorded the participant's response via the computer keyboard). Instructions to participants emphasized both speed and accuracy. Instructions also stressed the importance of exploring the hand-held object with both hands throughout the trial to ensure rich haptic shape acquisition. Participants were told that the held object was no more likely to be a target object than a distractor object. The experimenter's entry of the participant's response triggered the next trial.

A test session consisted of 64 trials in total, evenly divided between the four conditions, but in different random orders for each participant. Each of the 16 haptic objects was presented once in three of the priming conditions (*identity*, *category*, and *unrelated*), and the same objects appeared as visual targets once in each of the conditions (including *neutral*). This counterbalancing ensured that any difference in accuracy across conditions was not confounded with different object pairs. It also ensured that object familiarity was a constant across conditions. The experimenter varied the orientations of the haptic objects, placing them in the hands of the participant in one of four orientations, unrelated to the visual images, which the experimenter could not see. The orientations of each visual object were also varied randomly between four different views. Across the 16 participants in the experiment, each object was presented 4 times in each of the 4 possible orientations. Prior to testing, participants were given four practice trials identical to the experimental trials in which the data were not recorded. A different set of objects was used for the practice trials.

Baseline Biases. The baseline probability that the haptic prime had the same label as the visual target was 50% throughout the experiment. However, in one of the four conditions (i.e., neutral) the haptic prime was an object that participants knew was not a possible visual target. Thus, when the prime was a potential visual target (i.e., in the identity, category, and unrelated conditions) there was a 67% chance that the object in the hand had the same label as the visual target. In order to assess the extent of any bias stemming from the hand-held object, we evaluated participants' responses against the baseline likelihood of responding based solely on the label of the object being held. The unrelated condition provided a second opportunity to assess the impact of biases that might arise from the hand-held object. Since the held object in this condition did not correspond to the target object, we could compare the probability of responding with the held object's label and compare it to the baseline likelihood of guessing correctly based on this label (1/8 = 12.5%).

Results

The dependent measures were identification accuracy (i.e., percentage correct) and the image identification point (the visual resolution at the moment of identification). Figure 2A shows average accuracy for each of the four priming conditions. The main finding was that the small accuracy difference between the identity and category conditions was not significant, whereas the larger difference between the unrelated condition and both the identity and category conditions was significant. A repeated-measures ANOVA indicated a main effect of condition, $F(3, 45)=4.52$, $p< 0.001$, $MSE=68.85$. Fisher LSD tests indicated that accuracy in the identity condition was significantly higher than in the neutral, $t(45)=2.133$, $p<0.04$, and unrelated conditions, $t(45)=3.598$, $p<0.01$. Accuracy in the category condition was significantly higher than in the unrelated condition, $t(45)=2.268$, $p<0.04$. No other pair-wise comparisons were significant.

Figure 2B shows the mean resolution of the visual image at the time a correct visual identification response was made. Lower values indicate better performance, as lower spatial frequency information was sufficient for identification. As for the previous measure, the small accuracy difference between the identity and category conditions was not significant, whereas the larger difference between the unrelated condition and both the identity and category conditions was significant. A repeated-measures ANOVA indicated significant differences among the conditions, $F(3,45)=9.83$, $p<0.001$, $MSE=2.368$. Fisher LSD tests indicated that images in the identity condition were identified at a lower image resolution than in the neutral, $t(45)=6.21$, $p<0.01$, or unrelated conditions, $t(45)=4.48$, $p<0.01$. Images in the category condition were also identified at a lower image resolution than those in the neutral, $t(45)=3.03$, $p<0.05$, or unrelated condition, $t(45)=3.48$, $p<0.02$. No other pair-wise comparisons were significant.

Given the significant priming in both the identity and category conditions over the unrelated condition, and the absence of a statistical difference between these two, a specific contribution from shared shape representations seems to be ruled out. However, the non-significant trend favoring identity over the category condition in both measures prompted us to examine the combined accuracy and image point data more closely. For example, if some participants adopted a conservative strategy, withholding responses until they were completely certain of the visual target, such a strategy might mask the influence of haptic information at earlier stages in the identification process, when visual uncertainty was greater (Ernst and Banks 2002; Helbig and Ernst 2007). If this is the case, one would predict that it would be most likely to find evidence of shape priming for responses made at the lowest image resolution, where visual information was least certain. Greater uncertainty of visual information should induce greater reliance on haptics, meaning that both the shape and the semantic signals from the haptic modality might contribute to priming when access to visual confirmation was impaired or slowed down. In the presence of a stronger visual signal, when responses are made at higher levels of image resolution, shape information from the haptic modality might be ignored because it is inherently slower than semantic information, or because participants deem it unnecessary at a strategic level, given ready access to the visual objects' label.

To test this possibility, we examined the identification accuracy data from Experiment 1 separately at three levels of image resolution, as shown in Figure 2C. The data were divided into terciles, corresponding to three levels of image resolution, and therefore visual uncertainty, across all participants. In this analysis, a repeated measures ANOVA indicated a significant condition x tercile interaction ($F(3, 45) = 2.35$, $p < .04$, $MSE = 0.018$). Fisher's LSD tests indicated that in the lowest image resolution tercile, where visual uncertainty was greatest, accuracy in the identity condition was significantly higher than in the category condition ($t(45) = 2.30$, $p < 0.03$), though not in the two upper terciles, where visual certainty was greater.

Baseline Biases. The semantic category of the haptic object matched the visual object on 67% of the trials on which participants held a task relevant object, and so we tested for a bias based on the haptic prime in two ways. First, when we examined accuracy in all conditions in which the held object could be a target (identity, category, unrelated), participants responded with the label of the held object 64% of the time, which was not higher than the optimal guessing rate of 67% (on 2/3 of trials the held object had the same label as the visual object). Second, we measured the extent to which responses matched the held object in the *unrelated* condition (where the haptic prime was a potential target, but did not correspond to the visual target in either identity or category). This occurred only 3.1% of the time, which was less than expected by chance (1/8 = 12.5%). For the analysis of accuracy as a function of different image resolutions (Figure 2C), we also tested separately for a bias at the lowest level of image resolution, where the tendency should be greatest. The results showed that the responses in the *unrelated* condition still matched the held object in the *unrelated* condition less than the chance level of 12.5% (obtained mean = 7.5%). In the identity, category and unrelated conditions, participants responded with the label of the held object 71.9%, which was not significantly different from 67%. These analyses gave us confidence that strategic guessing on the part of participants did not contaminate visual identification accuracy. Note, too that if this strategy was employed it would still not account for differences between category and identity conditions, which are used to assess shape priming. This is because in these two conditions, objects share the same response and therefore the same bias.

Discussion

The main finding of Experiment 1 was that semantic congruency between seen and felt objects increased accuracy in naming a visual target object and decreased the amount of information needed to correctly identify it. These findings point to a robust semantic priming effect. Evidence of shape priming was not as clear. When responses made at all levels of image resolution were considered, the small benefit of identity versus category priming was not statistically reliable. Taken at face value, this finding supports critics saying that semantic priming may have played a larger role than acknowledged by Reales and Ballesteros (1999) (Craddock and Lawson 2008; Ernst et al. 2007; James et al. 2002).

However, a closer examination of the data, considering responses made at the lowest levels of image resolution, pointed to evidence for shape-based priming in addition to semantic priming. At the lowest tercile of image resolution, identity priming was significantly larger than category priming. This finding raises the possibility that shape priming may be more likely when there is greater uncertainty about the visual signal. This is consistent with the theory that information from different sensory modalities is integrated in a statistically optimal fashion (Ernst and Banks 2002). To illustrate, participants in a study by Helbig and Ernst (2007) were asked to make shape judgments of elliptical objects that were presented simultaneously, but varied slightly in their shape. When the amount of noise in the visual display was systematically varied, the results supported visual dominance at low levels of noise (haptic shape was ignored) and integration at higher levels of noise (haptic shape was incorporated). By analogy, it is possible that the open-ended procedure of Experiment 1 favored visual dominance and therefore only semantic priming from the held object. The reliable visual signal that was always available if the participant simply waited long enough, may have limited access to the haptic shape information that was more salient at lower levels of visual resolution. We test this hypothesis in our next experiment.

Experiment 2

This experiment was identical to Experiment 1, with the exception that we stopped the progressive revelation of the visual target when it reached 50% of full resolution and asked participants to respond as best they could. We chose this point in the sequence because it corresponded to the upper bound of image resolution in the lowest tercile of Experiment 1, where shape seemed to be influencing priming. Based on optimal integration theory (Ernst and Banks 2002), we anticipated that reducing the reliability of the visual signal would increase the influence of haptic-shape information on visual target identification.

Methods

Sixteen participants (7 female), between 20 and 31 years of age from the same population as Experiment 1 volunteered for the experiment. All were right-handed and had normal or corrected-to-normal visual acuity, and gave written informed consent. The procedures were followed here as in Experiment 1, with the exception that participants were required to make a response when the image sequence was at approximately 50% of its native resolution (image filter at the frequency band of 67-132 cycles/pixel), which was reached in 7 sec. Participants did not need to use the foot pedal in this experiment because of the fixed length of the image sequence.

Results

Figure 3 shows mean accuracy (percentage correct) of visual identification in the four priming conditions. There were significant differences in accuracy between identity versus category conditions, and between category versus unrelated conditions. A repeated measures ANOVA indicated significant differences among the conditions, $F(3, 45)=14.38$, $p< 0.001$, $MSE=97.2$.

Fishers LSD tests showed a significant benefit in accuracy in the identity condition over the category, $t(45)=3.36$, $p<0.01$, neutral, $t(45)=2.13$, $p<0.05$, and unrelated conditions, $t(45)=3.60$, $p<0.01$. Accuracy in the category condition was also significantly higher than in the unrelated condition, $t(45)=2.27$, $p<0.05$. No other pair-wise comparisons were significantly different. These findings indicate a clear benefit of shared shape in the haptic priming of a visual target (identity vs. category) over and above the benefit of semantics (category vs. unrelated).

We also compared the shape priming effect across Experiments 1 and 2. A mixed ANOVA comparing the between-groups factor of experiment and the repeated measures factor of shape priming (identity versus category conditions) indicated a significant interaction, $F(1, 30)=5.90$, $p<.03$, $MSE=41.29$, indicating that the 11.7% shape priming effect in Experiment 2 was significantly larger than the 3.9% effect in Experiment 1.

Power calculations based on G*Power's t-tests for dependent sample (Faul et al., 2007) with the estimates of population variance based on these data, indicated that the design of both Experiments 1 and 2 had the statistical power to detect differences between conditions of 5-6% at least 80% of the time.

Baseline Biases. As in Experiment 1, we again tested the possibility that the priming effects were influenced by biases based on the held priming object. Responses matched the held object in the unrelated condition 7.8% of the time, which was less than what would be expected by chance (1/8 = 12.5%). When we examined the responses across all the conditions in which the held object could be a target (identity, category, unrelated), the responses matched the label of the held object 59% of the time, which was not higher than the optimal guessing rate of 67%. These analyses imply that the haptic priming effects were not contaminated by strategic guessing based on the held object. Note that even if such contamination existed, it would not account for the difference between identity and category conditions, because both objects in those conditions share the same label and therefore the same bias.

Discussion

In Experiment 2 we reduced the certainty of visual information to address whether visual identification would benefit from haptic shape priming at reduced levels of visual precision. The results showed clear evidence of shape-based priming, as seen in the significant accuracy advantage in the identity priming condition over the category priming condition. A second finding in Experiment 2 was that visual identification accuracy in the category condition was significantly greater than in the unrelated condition. This replicated Experiment 1 in showing evidence for semantic priming over and above shape priming. Moreover, accuracy in the neutral condition lay midway between these two points in both experiments. This implies that haptic information not only benefits visual identification when it is consistent in shape, but that it interferes with visual identification when it is inconsistent. This disruption can be considered goal-oriented, because it occurs when the held object is also a potential visual target. Such interference is consistent with Bayesian inference models that propose that our perceptual experience is derived from combining prior knowledge and assumptions with current sensory evidence from different modalities (Ernst and Banks 2002).

General Discussion

The present study isolated the contributions of shared shape representations from semantic associations in haptic priming of visual object identification. Previous studies have not conclusively distinguished the relative contribution of these two sources of priming (Reales and Ballesteros 1999; cf. Craddock and Lawson 2008; Ernst et al. 2007; James et al. 2002). The main finding of the two experiments reported here is that *both* shape and semantics mediate haptic-to-visual priming of familiar objects. A second finding is that the relative magnitude of haptic priming depends on the reliability of the visual information. When participants were free to accumulate as much visual information as they deemed necessary before making a response (Experiment 1), the primary benefits of haptic priming were almost entirely semantic. This meant that a haptic prime sharing a semantic label with a visual target primed the identification of the visual target even though the two objects were quite different in shape. Yet when participants made their visual identification responses based on lower resolution visual information – be it by choice (Experiment 1, lower tercile responses) or by the design of the experiment (Experiment 2) – evidence of shape-based priming was clear.

We interpret this dependency of shape priming on the quality of the visual signal from the perspective of optimal integration theory (Ernst and Banks 2002). This theory states that the weight of a sensory cue in multisensory decision-making is directly proportional to the cue's reliability relative to other available cues. Hence, reducing the reliability of information in one channel increases the relative weight of information acquired from the other channel. One notable consequence of such a weighting strategy is that in addition to benefiting perceptual decisions when the cue and target are congruent, it can also lead to errors when the cue and target are incongruent. In both experiments, participants' response accuracy was higher in the category condition than in the unrelated condition. Accuracy in a neutral condition, where participants held an object that had no potential to be the visual target (i.e., ping pong ball) lay between these two points. This pattern of results implies that haptic information influenced the visual identification task in a goal-directed fashion.

This overall pattern of results, showing both shape and semantic contributions, does not support previously made either-or claims for the source of haptic-visual priming. As such, the results speak against the idea that automatic semantic associations are the only link between felt and seen objects (Humphreys et al. 1999; Klatzky and Lederman 1999). They also speak against accounts of visual-haptic priming that refer only to amodal shape representations (Lewkowicz and Lickliter 1994; Walker-Andrews 1994), although we note that this theory was specifically tailored to account for cross-modal interactions involving *unfamiliar* objects. The present finding of a semantic contribution, in addition to one from shape, implies that these theoretical accounts are not universal, and so caution against generalizing results from unfamiliar to familiar objects.

The present findings are generally in agreement with dual-coding theories (Johnson et al. 1989; Paivio 1991), though they do not help distinguish between the details of these theories. For instance, Johnson et al.'s (1989) theory explicitly proposes that shape representations are fundamentally visual, which means that haptic shape information in the present study was first converted into a visual code before influencing visual identification. An alternative account is that shape representations in both modalities are mediated by amodal structural descriptions of shape, akin to volumetric primitives in Biederman's (1987) geon theory.

The finding that both shape and semantic information acquired through touch can influence visual object identification tempts us to speculate on the underlying brain mechanisms. Ploran et al.'s (2007) used the progressive revelation of visual images to investigate the sequential stages of brain network activation that support visual identification processes. These authors reported a dissociation between the brain regions involved in low-level visual processing (occipital regions), perceptual evidence accumulation (inferior temporal, frontal, and parietal regions) and identification decision (medial frontal cortex, anterior insula/frontal operculum, and thalamus). In light of the present evidence for both shape and semantic priming when the visual signal was less reliable (Experiment 2), but only semantic priming when the visual signal was strong (Experiment 1), it will be important in the future to test whether shape priming is linked to brain regions involved in perceptual accumulation and whether semantic priming is linked to brain regions involved in decision-making. Combining the present behavioral methodology with the neuroimaging methods of Ploran et al. (2007) could be used to test this hypothesis.

An issue that remains for future research concerns the relationship between the present results, obtained with simultaneous presentation of objects in both modalities, and the results of previous visual-haptic priming research that was largely based on the study-test paradigm (e.g., Craddock & Lawson, 2009; Lawson, 2009; Reales and Ballesteros 1999). In our view, simultaneous-presentation tasks have the advantage of indexing online perceptual processes more directly than study-test tasks, which are more reliant on memory. However, it is possible that simultaneous presentation is more prone to response bias, simply because the prime and target objects are at the center of the participant's awareness. An important future step will be to make direct comparisons between these classes of procedure.

Another fruitful avenue for further research will involve testing priming in the opposite direction, namely, using visual primes to influence haptic object identification. An important theoretical question is whether the present finding — relatively greater shape priming when visual certainty was reduced — occurred because vision is the dominant and more reliable sense, or because visual identification was the primary task of participants. This question can be answered by comparing haptic-to-visual priming (as at present) with a visual-to-haptic priming task. If the answer is that visual dominance is critical, then the same results will be found in visual-to-haptic priming, provided that visual input is sufficiently degraded. If, however, the critical factor is the primary task of the participant, then the present results pattern will occur in visual-to-haptic priming only when the haptic input is relatively more degraded.

In conclusion, these data contribute to the ongoing debate over the role of shape and semantics in visual-haptic interaction. Its methodological contribution is that the experimental design offers a way to index the separate contributions of these factors in cross-modal priming of familiar objects. The theoretical contribution stems directly from the results, which indicate that haptic shape priming is more likely to occur when the reliability of vision is reduced relative to touch. As such, these results imply that the benefits of touch in your search for an item in your backpack, will indeed be larger at dusk than they are during the day.

Acknowledgments

This work was supported by a PhD scholarship to author AP from the Portuguese Fundação para a Ciência e Tecnologia (SFRH/BD/76087/2011), an Exchange Fellowship to author AAB from the Dr. Michael Quinn Memorial Fund, Department of Psychology, University of British Columbia, an NSERC (Canada) Discovery Grant to author JTE, and grants to author SS-F from the Spanish Ministry of Science and Innovation (PSI2010-15426 and Consolider INGENIO CSD2007-00012 grants, theComissionat per a Universitats i Recerca del DIUE-Generalitat de Catalunya (SRG2009-092), and the European Research Council (StG-2010 263145).

References

Amedi A, Malach R, Hendler T, Peled S, Zohary E (2001) Visuo-haptic object-related activation in the ventral visual pathway. Nature Neuroscience 4:324–30. doi:10.1038/85201

Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ (2005) Functional imaging of human crossmodal identification and object recognition. Experimental Brain Research 166:559–571. doi:10.1007/s00221-005-2396-5

Ballesteros S, González M, Mayas J, García-Rodríguez B, Reales JM (2009) Cross-modal repetition priming in young and old adults, European Journal of Cognitive Psychology 21(2/3): 366-387. doi: 10.1080/09541440802311956

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. Psychological Review 94(2):115-147.

Biederman I, Cooper EE (1991) Evidence for complete translational and reflectional invariance in visual object priming. Perception 20:585–593.

Biederman I, Cooper EE (2009) Biederman and Cooper's 1991 paper Translational and reflectional priming invariance: a retrospective. Perception 38:809–826. doi:10.1068/ldmk-bie

Bushnell EW, Baxt C (1999) Children's haptic and cross-modal recognition with familiar and unfamiliar objects. Journal of Experimental Psychology: Human Perception and Performance 25:1867–1881.

Carr TH, McCauley C, Sperber RD, Parmelee CM (1982) Words, pictures, and priming: on semantic activation, conscious identification, and the automaticity of information processing. Journal of Experimental Psychology: Human Perception and Performance 8:757–77. doi:10.1037/0096-1523.8.6.757

Craddock M (2008) Repetition priming and the haptic recognition of familiar and unfamiliar objects. Attention, Perception & Psychophysics 70:1350–1365. doi:10.3758/PP

Craddock M, Lawson R (2009) Size-sensitive perceptual representations underlie visual and haptic object recognition. PLoS ONE 4(11):e8009. doi:10.1371/journal.pone.0008009

Easton R, Greene AJ, 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 RD, Srinivas K, Greene AJ (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 MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends in Cognitive Sciences 8:162–9. doi:10.1016/j.tics.2004.02.002

Ernst MO, Lange C, Newell FN (2007) Multisensory recognition of actively explored objects. Canadian Journal of Experimental Psychology 61:242–253. doi:10.1037/cjep2007025

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–33. doi:10.1038/415429a

Faul F, Erdfelder E, Lang AG, Buchner A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods 39:175-191.

Felsen G, Dan Y (2005) A natural approach to studying vision. Nature Neuroscience 8:1643–1646. doi:10.1038/nn1608

Friese U, Supp GG, Hipp JF, Engel AK, Gruber T (2012) Oscillatory MEG gamma band activity dissociates perceptual and conceptual aspects of visual object processing: a combined repetition/conceptual priming study. NeuroImage 59:86171. doi:10.1016/j.neuroimage.2011.07.073

Glaser, WR (1992) Picture naming. Cognition 42:61–105. doi:10.1016/0010-0277(92)90040-O

Gottfried AW, Rose SA, Bridger WH (1977) Cross-modal transfer in human infants. Child Development 48:118–23.

Grefkes C, Weiss PH, Zilles K, Fink GR (2002) Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. Neuron 35:173–184.

Hadjikhani N, Roland PE (1998) Cross-modal transfer of information between the tactile and the visual representations in the human brain: A Positron Emission Tomographic study. J. Neuroscience 18:1072–1084

Helbig HB, Ernst MO (2007) Optimal integration of shape information from vision and touch. Experimental brain research 179:595–606. doi:10.1007/s00221-006-0814-y

Hillis AE, Caramazza A (1995) Cognitive and Neural Mechanisms Underlying Visual and Semantic Processing: Implications from “Optic Aphasia.” Journal of Cognitive Neuroscience 7:457–478. doi:10.1162/jocn.1995.7.4.457

Humphreys GW, Price CJ, Riddoch MJ (1999) From objects to names: a cognitive neuroscience approach. Psychological Research 62:118–30

Iordanescu L, Guzman-Martinez E, Grabowecky M, Suzuki S (2008) Characteristic sounds facilitate visual search. Psychonomic Bulletin & Review 15(3): 548-554

James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. Neuropsychologia 40:1706–1714. doi:10.1016/S0028-3932(02)00017-9

Johnson L, Paivio A, Clark J (1989) Spatial and verbal abilities in children’s cross-modal recognition: a dual-coding approach. Canadian Journal of Psychology 43:397–412.

Kahlaoui K, Baccino T, Joanette Y, Magnié MN (2007) Pictures and words: Priming and category effects in object processing. Current Psychology Letters Behavior Brain and Cognition 3:2–13.

Klatzky RL, Lederman SJ (1999) The haptic glance: A route to rapid object identification and manipulation. In: Gopher D, Koriat A (Eds) Attention and Performance XVII. The MIT Press, Cambridge, MA, pp. 164-196

Lacey S, Tal N, Amedi A, Sathian K (2009) A putative model of multisensory object representation. Brain Topography 21: 69–274. doi:10.1007/s10548-009-0087-4

Lacey S, Campbell C (2006) Mental representation in visual/haptic crossmodal memory: evidence from interference effects. Quarterly Journal of Experimental Psychology 59:361–76. doi:10.1080/17470210500173232

Lawson R. (2009) A comparison of the effects of depth rotation on visual and haptic three-dimensional object recognition. Journal of Experimental Psychology: Human Perception and Performance 35(4): 911-930. doi: 10.1037/a0015025

Lewkowicz DJ (1994) Development of intersensory perception in human infants. In: Lickliter DJ, Lewkowicz R (ed) The development of inter- sensory perception: Comparative perspectives. Hove UK, pp 165-203

Meltzoff AN, Borton RW (1979) Intermodal matching by human neonates. Nature 282:403–404

Neely JH (1977) Semantic Priming and Retrieval from Lexical Memory: Roles of Inhibitionless Spreading Activation and Limited-Capacity Attention. Journal of Experimental Psychology: General 106:226–254.

Newell FN, Ernst MO, Tjan BS, & Bulthoff HH (2001) Viewpoint Dependence in Visual and Haptic Object Recognition. Psychological Science 12:37–42. doi:10.1111/1467-9280.00307

Newell FN (2004) Cross-modal object recognition. In: Calvert G, Spence C, Stein BE (eds) The handbook of multisensory processes. MIT Press, Cambridge, MA, pp 123–139

Orgs G, Lange K, Dombrowski JH, Heil M (2006) Conceptual priming for environmental sounds and words: an ERP study. Brain and Cognition 62:267–72. doi:10.1016/j.bandc.2006.05.003

Paivio A (1991) Dual coding theory: Retrospect and current status. Canadian Journal of Psychology 45:255–287

Piai V, Roelofs A, Schriefers H (2011) Semantic interference in immediate and delayed naming and reading: Attention and task decisions. Journal of Memory and Language 64:404–423. doi:10.1016/j.jml.2011.01.004

Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME (2007) Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. The Journal of Neuroscience 27:11912–24. doi:10.1523/JNEUROSCI.3522-07.2007

Reales JM, Ballesteros S (1999) Implicit and explicit memory for visual and haptic objects: Cross-modal priming depends on structural descriptions. Journal of Experimental Psychology: Learning, Memory, and Cognition 25:644–663. doi:10.1037//0278-7393.25.3.644

Rentschler I, Juttner M, Osman E, Muller A, Caelli T (2004) Development of configural 3D object recognition. Behavioural Brain Research 149: 107–111. doi: 10.1016/S0166-4328(03)00194-3

Rose SA (1994) From hand to eye: Findings and issues in infant cross-modal transfer. In: Lickliter (ed) The development of intersensory perception: Comparative perspectives, Hove UK, pp 265–284.

Saito DN, Okada T, Morita Y, Yonekura Y, Sadato N (2003) Tactile-visual cross-modal shape matching: a functional MRI study. Cognitive Brain Research 17:14–25. doi:10.1016/S0926-6410(03)00076-4

Schacter DL, Dobbins IG, Schnyer DM (2004) Specificity of priming: a cognitive neuroscience perspective. Nature Reviews: Neuroscience 5:853–62. doi:10.1038/nrn1534

Stadthagen-Gonzalez H, Damian MF, Pérez MA, Bowers JS, Marín J (2009) Name-picture verification as a control measure for object naming: a task analysis and norms for a large set of pictures. Quarterly Journal of Experimental Psychology 62:1581-1597. doi:10.1080/17470210802511139

Tipper SP, Driver J (1988) Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. Memory & Cognition 16:64–70. doi:10.3758/BF03197746

Walker-Andrews AS (1994) Taxonomy for intermodal relations. In: Lewkowicz DJ, Lickliter R (eds) The development of intersensory perception: Comparative perspectives, Hove UK, pp 39–56

Welch R, Warren D (1986) Handbook of perception and human performance. Volume 1. Wiley-Interscience, New York.

Figure Captions

Figure 1. (A) The four visual-haptic priming conditions in Experiments 1 and 2, illustrated for the visual target "light bulb." (B) The total set of 16 familiar objects used as haptic primes and visual targets in Experiments 1 and 2.

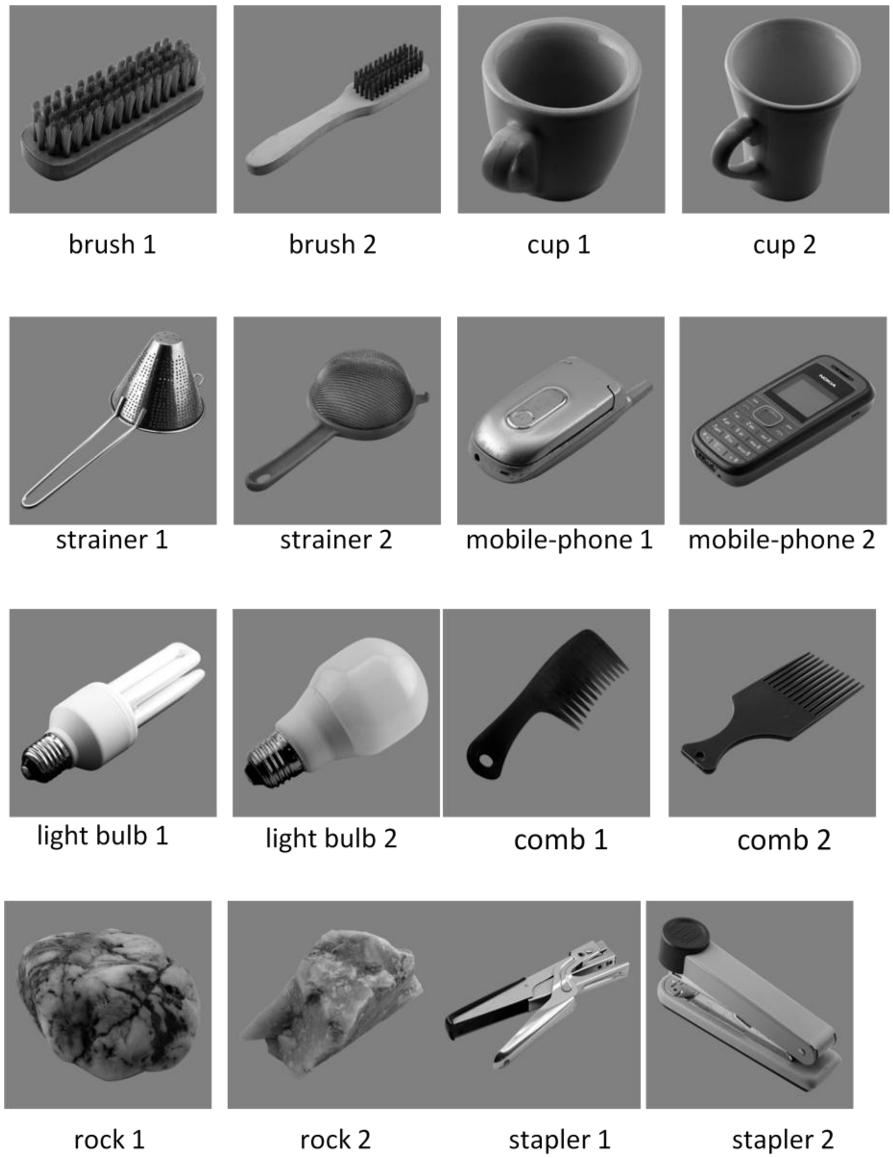


Figure 2. Results in the four haptic priming conditions of Experiment 1. (A) Mean visual identification accuracy. (B) Mean image resolution at the moment of correct visual identification (expressed as a percentage of full image resolution). (C) Mean visual identification accuracy after the data have been subdivided into three equal bins (terciles) of threshold image resolution. Error bars represent one standard error of the mean.

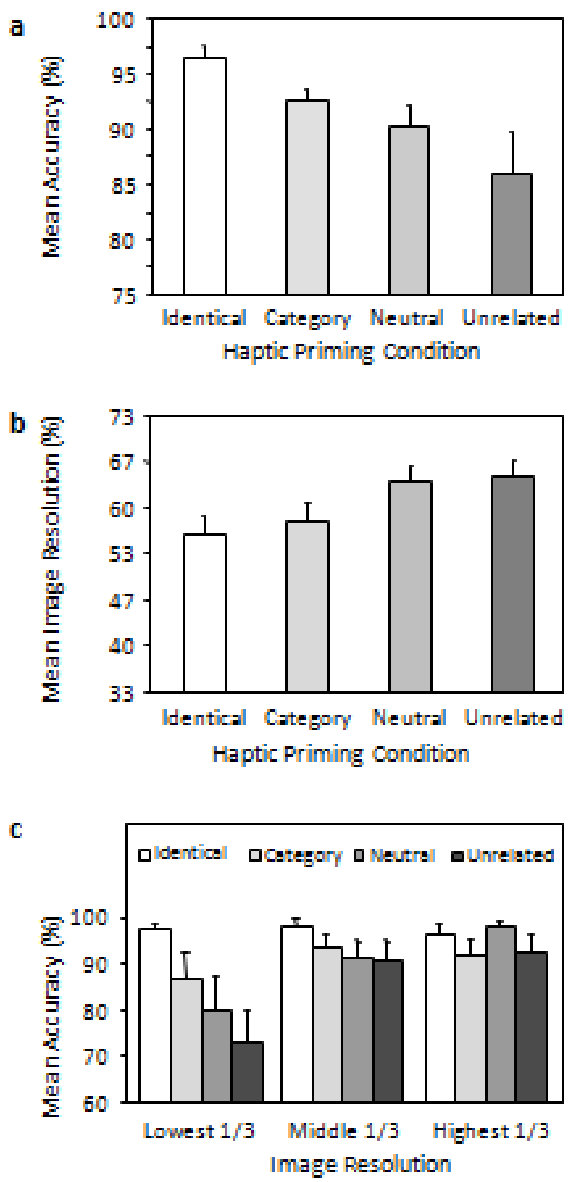
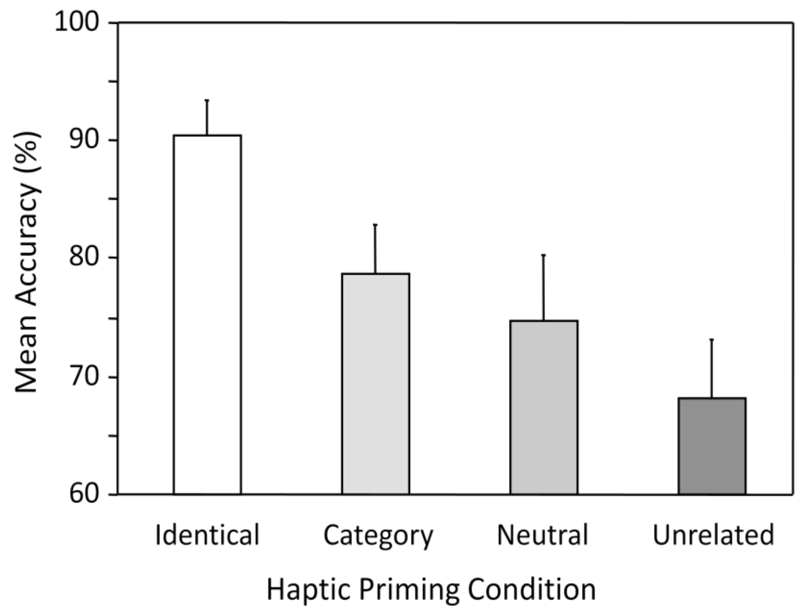


Figure 3. Mean visual identification accuracy in Experiment 2, where image resolution never exceeded 50%. Error bars represent one standard error of the mean.



Online Resources

Online Resource 1 An example of the progressive revelation of a target object used in Experiment 1 (mp4 file).

Online Resource 2 Photos of the Experimental set-up used in Experiment 1 and 2. (A) View of the participant (left) and experimenter (right). (B) Participant's view of the visual target sequence. (C). Experimenter's view of the haptic prime being explored by the participant.