Haptic guidance of overt visual attention

Alexandra List · Lucica Iordanescu ·
Marcia Grabowecky · Satoru Suzuki

© The Psychonomic Society, Inc. 2014

Abstract Research has shown that information accessed from one sensory modality can influence perceptual and attentional processes in another modality. Here, we demonstrated a novel crossmodal influence of haptic-shape information on visual attention. Participants visually searched for a target object (e.g., an orange) presented among distractor objects, fixating the target as quickly as possible. While searching for the target, participants held (never viewed and out of sight) an item of a specific shape in their hands. In two experiments, we demonstrated that the time for the eyes to reach a target—a measure of overt visual attention—was reduced when the shape of the held item (e.g., a sphere) was consistent with the shape of the visual target (e.g., an orange), relative to when the held shape was unrelated to the target (e.g., a hockey puck) or when no shape was held. This haptic-to-visual facilitation occurred despite the fact that the held shapes were not predictive of the visual targets' shapes, suggesting that the crossmodal influence occurred automatically, reflecting shape-specific haptic guidance of overt visual attention.

Keywords Multisensory processing · Visual search · Attention

Visual attention is known to be influenced by signals from other sensory modalities. In the domain of spatial attention, auditory, tactile, and visual cues produce spatially specific attentional facilitation or inhibition of visual processing (e.g., Driver & Spence, 1998). Converging evidence from patients with spatial attention deficits, such as neglect or extinction, has shown that auditory or tactile signals presented in one hemifield compete with visual signals in the opposite hemifield for drawing visual attention, as occurs with competing bilateral visual signals (e.g., Brozzoli, Demattè, Pavani, Frassinetti, & Farnè, 2006). In addition to crossmodal spatial interactions, feature- and object-based crossmodal influences of audition on visual attention have been reported (Guzman-Martinez, Ortega, Grabowecky, Mossbridge, & Suzuki, 2012; Iordanescu, Grabowecky, Franconeri, Theeuwes, & Suzuki, 2010; Iordanescu, Grabowecky, & Suzuki, 2011; Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki, 2008). For example, the localization of visual target (e.g., a cat) during search is facilitated by hearing its characteristic sound (e.g., “meow”), even when the sound neither predicts the target’s identity (which was unrelated to targets on 2/3 of the trials) nor provides any spatial information. This type of object-based crossmodal effect introduces the possibility that other nonspatial crossmodal interactions can influence the allocation of visual attention.

In the present study, we examined whether or not haptic-shape information influences visual attention, as illustrated by the following scenario. Imagine searching for an AA battery on a cluttered desk. It will not emit a beep (an auditory spatial cue), nor can you produce or play a characteristic battery sound (an auditory identity cue). What then might help your search? Here, we tested whether locating a rogue cylindrical battery would be facilitated by holding a similarly cylindrical water bottle. This is a new direction in haptic–visual research, addressing how haptic-shape information might influence visual exploratory behavior.

A. List · L. Iordanescu · M. Grabowecky · S. Suzuki
Department of Psychology, Northwestern University, Evanston,
IL, USA

M. Grabowecky · S. Suzuki
Interdepartmental Neuroscience Program, Northwestern University,
Evanston, IL, USA

A. List (✉)
Department of Psychology, Hamilton College,
198 College Hill Road, Clinton, NY 13323, USA
e-mail: alist@hamilton.edu

Previous research on haptic–visual interactions has focused on the recognition of three-dimensional (3-D) shapes. Early findings suggested visual dominance over haptics in shape identification (Miller, 1972; Power, 1981; Rock & Victor, 1964; Welch & Warren, 1980). Such findings of visual dominance over haptics may be due to the greater spatial resolution of the visual modality (e.g., Loomis, Klatzky, & Lederman, 1991). Consistent with this idea, when visual information is sufficiently distorted or restricted, haptic information can also influence visual shape identification (e.g., Heller, 1983). Furthermore, Ernst and colleagues (e.g., Ernst & Banks, 2002; Ernst & Bühlhoff, 2004) demonstrated that unimodal visual and haptic signals are weighted according to their reliability and are then combined, thereby statistically optimizing their contributions to the multimodal perceptual interpretation. However, recent results have suggested that, although the visual modality tends to be superior to the haptic modality for 3-D shape discrimination when people are allowed to view or feel shapes only from a restricted perspective, the haptic modality is just as effective as (or even superior to) the visual modality when people are allowed to freely view or manipulate objects (e.g., Norman et al., 2006; Norman et al., 2012).

Particularly relevant to the present study, 3-D shape discrimination across visual and haptic modalities (i.e., comparing a felt shape with a viewed shape, or vice versa) can be as accurate as unimodal shape discrimination (e.g., Norman, Clayton, Norman, & Crabtree, 2008; Norman et al., 2006). Furthermore, 3-D shape identification across visual and haptic modalities (i.e., learning shapes in the visual modality and identifying them in the haptic modality, or vice versa) can be relatively viewpoint invariant and nearly as accurate as unimodal shape identification (Lacey, Peters, & Sathian, 2007). A study by Gaißert, Wallraven, and Bühlhoff (2010) further suggests that visual and haptic representations of 3-D shape share similar topological structure. Consistent with the behavioral evidence of visual–haptic correspondences in the perception of 3-D shapes, neuroimaging research suggests that a portion of lateral occipital complex, known as the lateral–occipital tactile–visual area (LOtv), responds to shape information, whether it is visually or haptically presented (e.g., Amedi, Jacobson, Hendler, Malach, & Zohary, 2002; Amedi, Malach, Hendler, Peled, & Zohary, 2001; James et al., 2002; James, Kim, & Fisher, 2007). It is thus plausible that haptically experiencing 3-D shape information might guide visual attention to similarly shaped objects, assuming that multisensory ventral areas such as LOtv, which process both visual and haptic shape information, interact with dorsal attention areas (e.g., IPS and FEF) critical for visual search (e.g., Corbetta & Shulman, 2002; Eglin, Robertson, & Knight, 1991; Wardak, Ibos, Duhamel, & Olivier, 2006). Although extensive research has demonstrated the contributions of visual and haptic modalities to 3-D shape perception,

to our knowledge, it remains unknown whether and how haptic–shape information might influence visual attention.

To address this issue, we tracked participants' eye movements during visual search (Fig. 1, top) while they held different (unseen) items in their hands. The held items and visual targets varied in their shape similarity (Fig. 1, bottom row). The shape of the held item could match the shape of the visual target (target-consistent), one of the visual distractors (distractor-consistent), or none of the visual objects (unrelated). These haptic–visual conditions were randomly intermixed with trials during which participants held no item (visual-only). The visual search target was auditorily announced on each trial, and participants were encouraged to look at the target's center as quickly as possible. We monitored how rapidly their eyes reached the target, providing a measure of how rapidly their overt visual attention reached the target. Using this approach, we demonstrated that the perception of haptic shape from holding an unseen 3-D item guides overt visual attention to a similar shape in the environment.

Fig. 1 Example of a visual search array (top). The search target was announced auditorily prior to the onset of the search array. An example for each of the three haptic–visual conditions is shown (bottom) for a trial in which the visual search target is a leash. The held shape either matched the target shape (target-consistent, left bottom), matched the shape of a distractor diametrically opposite from the target (distractor-consistent, middle bottom), or did not match the shape of any of the images in the visual search array (unrelated, right bottom). The visual images shown are only illustrative of those used in the experiment

Experiment 1: Visual search with a concurrent haptic memory task

Method

Participants

Ten members of the Northwestern University community (five women, five men; all right-handed; $M = 23$ years of age) gave informed consent to participate. All reported normal sensation in their hands, normal or corrected-to-normal vision, and normal hearing.

Apparatus

Eye movements were recorded with an EyeLink 2000 tower-mount eyetracker, sampling the left eye's position at a temporal resolution of 1000 Hz and a spatial resolution of 0.25° . A centered 9-point rectangular array (a 3×3 grid, spanning $18^\circ \times 24^\circ$) was used for calibration. Saccades were detected using the standard EyeLink thresholding method, based on eye position shift (0.1°), saccade velocity ($30^\circ/\text{s}$), and saccade acceleration ($8000^\circ/\text{s}^2$), in conjunction with the general algorithm described in Stampe (1993). A 20-in. P225f Viewsonic CRT monitor was used to present visual stimuli with a refresh rate of 75 Hz and spatial resolution of $1,024 \times 768$ pixels. The viewing distance was 84 cm. Stimulus presentation and response recording were carried out with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and the EyeLink Toolbox (Cornelissen, Peters, & Palmer, 2002) in the MATLAB environment (www.mathworks.com).

Stimuli

Haptic stimuli Twelve graspable items (all wooden blocks except for the sphere, which was plastic) were used as haptic stimuli. The items were the following shapes: arch, cone, cube, high cylinder, high parallelepiped, low cylinder, low parallelepiped, low square, sphere, squiggle, star, and triangle (see Table 1 for the dimensions). These shape-category names were never mentioned to participants.

Visual stimuli The visual images were grayscale photographs of real-world objects that had shapes similar to those of the 12 graspable items. Prior research has shown that the assumed real-world sizes of objects are automatically accessed from visual images (e.g., Ittelson, 1951; Konkle & Oliva, 2011, 2012). Thus, for each haptic shape, we included three visual images depicting the shape in relatively small, intermediate, and large assumed real-world sizes. For instance, for the haptic spherical shape, we included images of an orange (smallest), a soccer ball (intermediate), and a planet (largest); see Table 1 for a complete list of the visual images. We aimed

to determine whether haptic–visual effects were specific to visual images of graspable objects or would generalize to visual images of any assumed size. Each of the 36 visual images ($12 \text{ shapes} \times 3 \text{ assumed sizes}$) subtended approximately $5.5^\circ \times 5.5^\circ$. Each search array consisted of eight images (two in each quadrant) presented on a uniform gray background (e.g., Fig. 1). The images were equally spaced and centered along an approximate iso-acuity ellipse ($19.2^\circ \times 14.4^\circ$, based on Virsu & Rovamo, 1979; Fig. 1). A white $0.8^\circ \times 0.8^\circ$ plus sign (line width = 0.14°) presented at the center of the display was used as the fixation marker.

Verbal labels Each visual stimulus had an associated auditory verbal label (Table 1). These verbal labels were validated in a pilot study with 18 participants. When asked to label each image, participants spontaneously generated half (on average) of the verbal labels exactly as listed in Table 1. With more lax criteria (e.g., *Jupiter* for *planet*, *puck* for *hockey puck*, or *house* for *building*), they generated 30/36 (on average) verbal labels, indicating that the labels were appropriate for announcing the search targets. The verbal labels had a duration of 460–1,300 ms ($M = 733 \text{ ms}$, $SD = 182$), and were presented auditorily (at $\sim 67 \text{ dB SPL}$) via a pair of speakers placed symmetrically just in front of and lateral to the monitor.

Procedure

Prior to starting the experiment, each participant was familiarized with all 36 visual images and their associated verbal labels (Table 1). Participants performed concurrent haptic and visual tasks: exploring an item out of sight with one or both hands to remember its 3-D shape while performing a visual search task.

At the start of each trial, the experimenter, who was unaware of the haptic–visual congruency condition, passed an item to and/or collected an item from the participant under a table while the participant's head was in a chinrest, to keep the item out of sight. The experimenter said one of three things: “switch” (the participant should hand over the last item and would be given a new item), “nothing” (the participant should hand over the last item, but no new item would be given), or “hold” (the participant should keep the item already in hand). As soon as the participant was holding the appropriate item (or nothing, depending on the condition), the experimenter pressed a key to start the visual search task.

Each visual search task began with the auditory verbal label (e.g., “hockey puck”) announcing the upcoming visual search target, and following a 1-s delay, the central fixation plus sign appeared. The participant was required to fixate within a central $2.5^\circ \times 2.5^\circ$ window for 500 ms, which triggered the appearance of a search array. The participant was instructed to look at the target item as quickly as possible and to fixate the center of the target image. We determined that the participant's

Table 1 Stimulus attributes

Haptic Shape	Dimensions (cm)	Visual Images and Verbal Labels (Ordered by the Assumed Size of the Depicted Objects)		
		Smallest	Intermediate	Largest
arch	$h = 2.8, w = 6, d = 3, r = 2$	handle	faucet	arch
cone	$h = 4.6, r = 1.7$	ice cream cone	party hat	teepee
cube	$h = w = d = 2.9$	dice	ice cube	building
high cylinder	$h = 6, r = 1.4$	soda	glass	tower
high parallelepiped	$h = w = 2.9, d = 6$	speaker	box	shipping container
low cylinder	$h = 1.5, r = 1.7$	hockey puck	drum	arena
low parallelepiped	$h = 6, w = 2.9, d = 1.4$	cell phone	chocolate	high-rise
low square	$h = w = 3, d = 1.5$	Post-it notes	gift	boxing ring
sphere	$r = 1.9$	orange	soccer ball	planet
squiggle	cross-section = $1 \times 0.8, w = 7.2, \text{peak-to-trough} = 1.4$	snake	leash	river
star	convex-vertex-to-center = 2.2, concave-vertex-to-center = 1.3, $d = 1.5$	badge	starfish	monument
triangle	equilateral side = 4, $d = 1.5$	triangle	pizza box	sailboat

For each haptic shape, its dimensions (h = height, w = width, d = depth, r = radius) and the shape-matched visual objects (identified by their verbal labels) of three different assumed real-world sizes are indicated

eyes reached the target when they landed within a $4.1^\circ \times 4.1^\circ$ “target window” centered on the target image. After 500 ms of target fixation, the display offset. A blank intertrial interval followed.

The eye-monitoring camera was adjusted and the eyetracker was calibrated to each participant at the beginning of the experiment. During the experiment, recalibrations and breaks were taken as needed. Saccadic search time was calculated as the time from the presentation of the stimulus array to the time when participants’ eyes initially reached the target window, whether or not participants maintained fixation at the location for 500 ms to indicate that they found the target. Thus, the saccadic search time indicates how long it took participants to move their overt attention to the target, independently of how long it took them subsequently to confirm the identity of the target.

A haptic memory test was given at the end of the experiment. Out of sight, the participant was handed each of the 12 items held during the experiment, as well as 12 new items, one at a time in a randomized order. The participant manually explored each item and responded whether it was new or had been held during the experiment.

Design

A new randomization of 144 trials was used for each participant. Each of the 36 images ($12 \text{ shapes} \times 3 \text{ assumed sizes}$) was presented as the target once in each of four haptic–visual conditions: target-consistent (the held shape matched the shape of the target image), distractor-consistent (the held shape matched the shape of a distractor image), unrelated (the held shape did not match the shape of any of the images),

and visual-only (no shape was held). The position of each target image was randomized, with the constraint that each target image was presented once in each quadrant.

Distractor images were randomly chosen (from the 35 images [36 images minus the target image]) and distributed in the array, with the following constraints. First, no more than one image depicting a specific haptic shape was presented within each visual search array. Second, some of the haptic shapes were similar to one another on the basis of geometric categories. In particular, the cube, low-square, low-parallelepiped and high-parallelepiped shapes were similar, in that they were all rectilinear; likewise, the low- and high-cylinder shapes were both cylindrical. When the target image depicted one of these shapes, none of the distractor images depicted shapes from the same geometric category. Third, when a distractor image was consistent with the shape of the held item, it was always presented diametrically opposite from the target image, so that if the haptic–visual interaction guided attention to the consistent distractor, attention would be directed away from the target.

Results and discussion

We analyzed saccadic search times (from the search array’s onset to a saccade reaching the target image) using a two-factor repeated measures analysis of variance (ANOVA) with haptic–visual condition (visual-only, target-consistent, distractor-consistent, and unrelated) and the assumed size of the visual target (smallest, intermediate, and largest) as the independent variables. We found a significant main effect of haptic–visual condition, $F(3, 27) = 10.159, p < .001, \eta_p^2 = .530$ (Fig. 2). Specifically, saccadic search times were

Fig. 2 Results of Experiment 1 (with a concurrent haptic memory task). Mean saccadic search times are shown for the four haptic-visual conditions: visual-only (holding no item), target-consistent (holding an item that matched the target shape), distractor-consistent (holding an item that matched the shape of a distractor), and unrelated (holding an item that did not match the shape of any of the visual images). The error bars represent ± 1 SE (adjusted for repeated measures comparisons). * $p < .05$, ** $p < .01$, *** $p < .001$

significantly faster when participants were holding an item with a target-consistent shape than when holding an item with a distractor-consistent shape, $t(9) = 8.649$, $p < .001$, $d = 2.735$, or holding an item with an unrelated shape, $t(9) = 3.544$, $p < .007$, $d = 1.120$. In addition, saccadic search times were significantly slower when holding an item with a distractor-consistent shape than when holding an item with an unrelated shape, $t(9) = 2.786$, $p < .022$, $d = 0.881$, or when holding no shape (visual-only condition), $t(9) = 2.960$, $p < .017$, $d = 0.936$. No other pairwise comparisons were significant, $|t|s < 1.6$, $ps > .16$. Thus, relative to holding a shape unrelated to any of the objects in the visual search array, holding a target-consistent shape sped saccadic search times, and holding a distractor-consistent shape slowed saccadic search times. However, holding a target-consistent shape did not significantly speed saccadic search times relative to holding nothing. This result was investigated in Experiment 2.

We also observed a significant main effect of assumed size, $F(2, 18) = 13.973$, $p < .001$, $\eta_p^2 = .608$. Specifically, saccadic search times were significantly slower for target images depicting the largest-size objects (e.g., house, monument, or teepee; $M = 814$ ms, $SE = 42$) relative to those depicting the intermediate-size objects (e.g., pizza box, soccer ball, or drum; $M = 696$ ms, $SE = 38$), $t(9) = 4.209$, $p < .003$, $d = 1.331$, and those depicting the smallest-size objects (e.g., badge, cell phone, or dice; $M = 665$ ms, $SE = 55$), $t(9) = 4.014$,

$p < .004$, $d = 1.269$; saccadic search times for the latter two assumed sizes did not differ significantly, $t(9) = 1.401$, $p > .19$, $d = 0.443$. This result may suggest that images depicting very large objects are more difficult to find than images depicting manipulable objects in a visual search context, but this interpretation is tentative, because the images depicting different assumed sizes in our set also differed in other aspects, such as texture, identifiable parts, and semantic category. Nevertheless, we found no significant interaction between the haptic-visual condition and assumed size, $F(6, 54) = 1.052$, $p > .40$, $\eta_p^2 = .105$, suggesting that the haptic-visual effect is shape-specific and independent of assumed size.

Performance on the haptic memory task, averaging 94.2 % hits and 2.4 % false alarms (indicating a high mean d' value of 3.549), confirmed that participants actively explored, attended to, and accurately remembered the shapes that they held.

The overall pattern of results, with significantly faster saccadic search times in the target-consistent condition and significantly slower saccadic search times in the distractor-consistent condition relative to the unrelated condition, supports the idea that holding a shape guides overt visual attention to similarly shaped objects. However, a few issues remain unresolved. First, holding a target-consistent shape did not speed search relative to holding nothing. One possibility is that having to memorize the held shape while performing the visual search task distracted participants' focus from the search task. If this were the case, holding a target-consistent shape should speed visual search relative to holding nothing, if participants do not have to simultaneously remember the held shape. Second, performing the memory task might have encouraged participants to verbalize the held shape, so that the obtained haptic-visual consistency effect could be mediated by a semantic-level interaction. If this were the case, the haptic-visual consistency effect should be substantially reduced or eliminated if participants do not have to remember the held shape. Third, maintaining the haptic-shape information in working memory for the later memory test may underlie the obtained influence of haptic-shape perception on visual attention. Soto and Humphreys (2007) reported that when maintaining, for example, "red square" in working memory for a posttrial test, visual attention was drawn to a red square during an intervening search task, even though the search task required a line orientation judgment that was irrelevant to the information in working memory. If maintaining shapes in working memory was critical for the obtained haptic-visual consistency effect, it should disappear if participants do not have to remember the held shapes.

To address these concerns, we conducted a follow-up experiment without a concurrent haptic memory task. Participants held no item, or were instructed to hold an item exclusively in either the right or the left hand. Our intent was to give participants the impression that we were testing the

effects of holding an item in the right versus the left hand on a visual task. Otherwise, this experiment was identical to Experiment 1.

Experiment 2: Visual search without a concurrent haptic memory task

Method

Participants

Fourteen members of the Northwestern University community (11 women, three men; 12 right-handed; $M = 23$ years of age) gave informed consent to participate. All reported normal sensation in their hands, normal or corrected-to-normal vision, and normal hearing.

Apparatus, stimuli, procedure, and design

The following modifications were made, relative to Experiment 1. The screen resolution was changed to $1,152 \times 864$ pixels, resulting in all visual spatial measurements being scaled down by a factor of .89. On alternating trials, participants held an item in either their right or their left hands. Half of the participants began with a “right-hand” trial, and the remaining participants began with a “left-hand” trial. The alternation continued through the randomly intermixed visual-only (holding no item) trials; for example, if a visual-only trial occurred on a “left-hand” trial, the next trial was a “right-hand” trial.

Results and discussion

As in Experiment 1, we analyzed saccadic search times using a two-factor repeated measures ANOVA with haptic–visual condition (visual-only, target-consistent, distractor-consistent, and unrelated) and the assumed size of the visual target (smallest, intermediate, and largest) as the independent variables. Again, we found a significant main effect of haptic–visual condition, $F(3, 39) = 2.905$, $p < .05$, $\eta_p^2 = .183$ (Fig. 3). Specifically, saccadic search times were significantly faster when holding an item with a target-consistent shape than when holding an item with a distractor-consistent shape, $t(13) = 2.165$, $p < .05$, $d = 0.579$, when holding an item with an unrelated shape, $t(13) = 3.050$, $p < .01$, $d = 0.815$, or when holding no item (visual-only condition), $t(13) = 2.343$, $p < .037$, $d = 0.626$. No other pairwise comparisons were significant, $|t|s < 1$. Thus, merely holding an item facilitated visual search when the visual target depicted the held shape.

Also as in Experiment 1, we observed a significant main effect of assumed size, $F(2, 26) = 20.263$, $p < .001$, $\eta_p^2 = .609$.

Fig. 3 Results of Experiment 2 (without a concurrent haptic memory task). Mean saccadic search times are shown for the four haptic–visual conditions: visual-only (holding no item), target-consistent (holding an item that matched the target shape), distractor-consistent (holding an item that matched the shape of a distractor), and unrelated (holding an item that did not match the shape of any of the visual images). The error bars represent ± 1 SE (adjusted for repeated measures comparisons). * $p < .05$, ** $p < .01$

Saccadic search times were slower for target images depicting the largest-size objects ($M = 947$ ms, $SE = 41$), relative to those depicting the intermediate-size objects ($M = 764$ ms, $SE = 30$), $t(13) = 5.943$, $p < .001$, $d = 1.588$, and those depicting the smallest-size objects ($M = 736$ ms, $SE = 41$), $t(13) = 4.836$, $p < .001$, $d = 1.293$; saccadic search times for the latter two sizes did not differ significantly, $|t| < 1$. Importantly, as in Experiment 1, we found no significant interaction between haptic–visual condition and assumed size, $F < 1$, suggesting that the haptic–visual interaction is shape-specific and independent of assumed size.

General discussion

To determine whether and how haptic-shape information influences visual attention during search, we gave participants unseen items to hold that were consistent with the shape of a visual target, a visual distractor, or none of the images in the visual array, and measured how long it took participants’ eyes to reach the target—a measure of how quickly participants’ overt visual attention reached the target. By design, haptic-shape information was antipredictive of the target; that is, the held shape was dissimilar to the visual target’s shape on two thirds of the haptic–visual trials. Nevertheless, haptic-shape information produced a systematic facilitative influence on visual exploratory behavior.

In Experiment 1, we instructed participants to remember the shape of each item they held (for a later haptic memory test), because we thought that actively encoding the haptic shape might be crucial for influencing overt visual attention in a shape-specific manner. Saccadic search times were faster when the held shape matched the target shape (target-consistent), and slower when the held shape matched the shape of a distractor shape (distractor-consistent), relative to when the held shape did not match the shape of any of the objects in the search array (unrelated). This result is consistent with the idea that holding a shape guides overt visual attention to objects of similar 3-D shapes.

The unrelated condition provided an appropriate control for the target-consistent and distractor-consistent conditions, because it was matched to the latter conditions in terms of any general dual-task effects of holding an item, exploring it, and trying to remember its haptic shape, except that the held item was dissimilar in shape to any of the visual objects. Nevertheless, the impact of the findings was limited, in that saccadic search times were not significantly faster when the held shape matched the target shape, relative to when no shape was held. Moreover, the use of the concurrent haptic memory task might have created unintended confounds. Therefore, we conducted a second experiment to control for several potential dual-task effects. First, we wanted to control for potential interference from the concurrent effort to remember the haptic shapes, which was absent when no shape was held. Second, we wanted to control for the possibility that the task of remembering a haptic shape might have encouraged verbalization of the shape (e.g., “a cube”), so that the results might have reflected a semantic rather than a haptic influence on visual search. Third, we wanted to control for the possibility that the effect might be driven by the act of maintaining the haptic shapes in working memory.

The second experiment was thus identical to the first, except that it included no haptic memory task; participants merely held an item (or held nothing) on each trial. Saccadic search times were significantly faster when the held shape matched the target shape relative to all other conditions, including when no shape was held in the visual-only condition (saccadic search times did not differ among the other conditions). Thus, the unique effect obtained in Experiment 1—namely, the significant slowing of saccadic search times in the distractor-consistent condition relative to the unrelated condition—could have been mediated by the active haptic exploration of the held item, the semantic encoding of the held shapes, and/or the maintenance of the held shapes in working memory, processes that were all encouraged by the concurrent haptic memory task. Nevertheless, in both experiments, we demonstrated that saccadic search times were faster in the haptic–visual shape-consistent condition than in the control conditions that were matched for any general effects of holding an item. Furthermore, the crossmodal effect

was independent of the assumed size of a visual target, confirming that the effect is driven by haptic–visual shape similarity. The results of Experiment 2 further demonstrate that merely holding an item matching the shape of a visual target, with no instruction or incentive to explore or remember the held shape, speeds saccadic search times relative to holding nothing. Thus, it is beneficial to hold in hand an item that matches the shape of a visual target during search.

Extensive behavioral and neuroimaging research has shown that 3-D shape representations in the visual and haptic modalities are closely associated and interactive (see the introduction). Our results suggest that this crossmodal association allows a haptically processed shape to highlight similarly shaped objects in the visual world, influencing visual exploratory behavior in a shape-specific manner. When searching for that rogue AA battery, hold on to your water bottle.

Author note This study was supported by NIH Grant Nos. R01 EY018197 and R01 EY021184, and by ARRA Supplement Grant No. R01 EY018197-02S1.

References

- Amedi, A., Jacobson, G., Hendler, T., Malach, R., & Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cerebral Cortex*, *12*, 1202–1212.
- Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience*, *4*, 324–330.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. doi:10.1163/156856897X00357
- Brozzoli, C., Demattè, M. L., Pavani, F., Frassinetti, F., & Farnè, A. (2006). Neglect and extinction: Within and between sensory modalities. *Restorative Neurology and Neuroscience*, *24*, 217–232.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*, 201–215.
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavioral Research Methods, Instruments, & Computers*, *34*, 613–617. doi:10.3758/BF03195489
- Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, *2*, 254–262.
- Eglin, M., Robertson, L. C., & Knight, R. T. (1991). Cortical substrates supporting visual search in humans. *Cerebral Cortex*, *1*, 262–272.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433. doi:10.1038/415429a
- Ernst, M. O., & Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*, 162–169. doi:10.1016/j.tics.2004.02.002
- Gaißert, N., Wallraven, C., & Bühlhoff, H. H. (2010). Visual and haptic perceptual spaces show high similarity in humans. *Journal of Vision*, *10*(11), 1–20. doi:10.1167/10.11.2
- Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology*, *22*, 383–388.

- Heller, M. A. (1983). Haptic dominance in form perception with blurred vision. *Perception*, 12, 607–613.
- Iordanescu, L., Grabowecky, M., Franconeri, S., Theeuwes, J., & Suzuki, S. (2010). Characteristic sounds make you look at target objects more quickly. *Attention, Perception, & Psychophysics*, 72, 1736–1741. doi:10.3758/APP.72.7.1736
- Iordanescu, L., Grabowecky, M., & Suzuki, S. (2011). Object-based auditory facilitation of visual search for pictures and words with frequent and rare targets. *Acta Psychologica*, 137, 252–259.
- Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, 15, 548–554. doi:10.3758/PBR.15.3.548
- Ittelson, W. H. (1951). Size as a cue to distance: Static localization. *American Journal of Psychology*, 64, 54–67.
- James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*, 40, 1706–1714.
- James, T. W., Kim, S., & Fisher, J. S. (2007). The neural basis of haptic object processing. *Canadian Journal of Experimental Psychology*, 61, 219–229. doi:10.1037/cjep.2007023
- Konkle, T., & Oliva, A. (2011). Canonical visual size for real-world objects. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 23–37. doi:10.1037/a0020413
- Konkle, T., & Oliva, A. (2012). A familiar-size Stroop effect: Real-world size is an automatic property of object representation. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 561–569. doi:10.1037/a0028294
- Lacey, S., Peters, A., & Sathian, K. (2007). Cross-modal object recognition is viewpoint-independent. *PLoS ONE*, 2, e890. doi:10.1371/journal.pone.0000890
- Loomis, J. M., Klatzky, R. L., & Lederman, S. J. (1991). Similarity of tactual and visual picture recognition with limited field of view. *Perception*, 20, 167–177.
- Miller, E. A. (1972). Interaction of vision and touch in conflict and nonconflict form perception tasks. *Journal of Experimental Psychology*, 96, 114–123.
- Norman, J. F., Clayton, A. M., Norman, H. F., & Crabtree, C. E. (2008). Learning to perceive differences in solid shape through vision and touch. *Perception*, 37, 185–196. doi:10.1068/p5679
- Norman, J. F., Crabtree, C. E., Norman, H. F., Moncrief, B. K., Herrmann, M., & Kapley, N. (2006). Aging and the visual, haptic, and cross-modal perception of natural object shape. *Perception*, 35, 1383–1395. doi:10.1068/p5504
- Norman, J. F., Phillips, F., Holmin, J. S., Norman, H. F., Beers, A. M., Boswell, A. M., ... Ronning, C. (2012). Solid shape discrimination from vision and haptics: Natural objects (*Capsicum annum*) and Gibson's "feelies." *Experimental Brain Research*, 222, 321–332.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. doi:10.1163/156856897X00366
- Power, R. P. (1981). The dominance of touch by vision: Occurs with familiar objects. *Perception*, 10, 29–33.
- Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, 143, 594–596.
- Soto, D., & Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 730–737. doi:10.1037/0096-1523.33.3.730
- Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers*, 25, 137–142.
- Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research*, 37, 475–494.
- Wardak, C., Ibos, G., Duhamel, J. R., & Olivier, E. (2006). Contribution of the monkey frontal eye field to covert visual attention. *Journal of Neuroscience*, 26, 4228–4235.
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638–667. doi:10.1037/0033-2909.88.3.638