

Cross-modal associations between vision, touch, and audition influence visual search through top-down attention, not bottom-up capture

Emily Orchard-Mills · David Alais · Erik Van der Burg

Published online: 24 August 2013
© Psychonomic Society, Inc. 2013

Abstract Recently, Guzman-Martinez, Ortega, Grabowecky, Mossbridge, and Suzuki (Current Biology : CB, 22(5), 383–388, 2012) reported that observers could systematically match auditory amplitude modulations and tactile amplitude modulations to visual spatial frequencies, proposing that these cross-modal matches produced automatic attentional effects. Using a series of visual search tasks, we investigated whether informative auditory, tactile, or bimodal cues can guide attention toward a visual Gabor of matched spatial frequency (among others with different spatial frequencies). These cues improved visual search for some but not all frequencies. Auditory cues improved search only for the lowest and highest spatial frequencies, whereas tactile cues were more effective and frequency specific, although less effective than visual cues. Importantly, although tactile cues could produce efficient search when informative, they had no effect when uninformative. This suggests that cross-modal frequency matching occurs at a cognitive rather than sensory level and, therefore, influences visual search through voluntary, goal-directed behavior, rather than automatic attentional capture.

Keywords Visual search · Attention · Cross-modal · Visual selection

Introduction

Visual search underpins many everyday tasks, such as looking for a book on a shelf or a friend in a crowd. Search sensitivity and specificity can also be trained and optimized for tasks such as baggage screening and radiographic diagnosis (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004; Nodine & Kundel,

1987; Wang, Lin, & Drury, 1997). As such, visual search is an ecologically relevant paradigm that gives insight into visual processing under the influence of strategic goals. In the experimental visual search paradigm, manipulation of the number of elements in the display can be used to assess search efficiency; search times generally increase by 150–300 ms per item if eye movements are required (Wolfe & Horowitz, 2004). However, if a feature captures attention, such increases may be nullified, since the target is rapidly separated from the surrounding elements without serial self-terminating scanning.

There are a number of attributes shown to allow efficient visual search, including color, motion, orientation, and size (for a discussion of these and other attributes, see a review by Wolfe & Horowitz, 2004). Targets that are unique along one of these dimensions are termed singletons and create a pop-out effect. The clearest example is a red target displayed among green nontarget items. Analogous to this effect in vision, search within the auditory domain is also influenced by features able to capture attention. One such example is the “cocktail-party” effect (Cherry, 1953; see Bronkhorst, 2000, for a review). Words of importance—for instance, one’s name—have been shown to pop out of an unattended conversation. Similarly, in touch, a movable item has been shown to create a pop-out effect in a display of static items (van Polanen, Tiest, & Kappers, 2012), as has a rough surface in a display of fine-textured items (Plaisier, Bergmann Tiest, & Kappers, 2008). Auditory and tactile cues have also been shown to influence visual search. This has been shown when the auditory or tactile cue was spatially informative (Bolia, D’Angelo, & McKinley, 1999; Jones, Gray, Spence, & Tan, 2008; Rudmann & Strybel, 1999), when the auditory or tactile cue was temporally synchronous with a change in color of the target (Ngo & Spence, 2010; Van der Burg, Cass, Olivers, Theeuwes, & Alais, 2010; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008b; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2009; Zannoli, Cass, Mamassian, & Alais, 2012) and when the auditory cue was semantically congruent

E. Orchard-Mills (✉) · D. Alais · E. Van der Burg
School of Psychology, University of Sydney, Sydney,
New South Wales 2006, Australia
e-mail: emily.orchardmills@sydney.edu.au

with the target object (Iordanescu, Grabowecky, Franconeri, Theeuwes, & Suzuki, 2010; Iordanescu, Gravowecy, & Suzuki, 2011; Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki, 2008).

A recent study by Guzman-Martinez, Ortega, Grabowecky, Mossbridge, and Suzuki (2012) reported that participants matched auditory amplitude modulation rate with visual spatial frequency with a linear relationship. The authors postulated that during manual exploration of surfaces, recurring multisensory inputs formed associations between visual spatial frequencies (i.e., the surface texture) and the modulation rate of auditory sounds that would be produced by touching them. They suggested that these links between auditory and visual frequencies resulted in automatic attentional capture such that when participants were presented with a modulating auditory sound, their attention was drawn to the matching visual spatial frequency. In order to support this suggestion, Guzman-Martinez et al. reported a number of experiments, one of which is important to the focus of this study. In this experiment (Guzman-Martinez et al., 2012, Experiment 4), participants were required to simultaneously monitor one very low (0.5 cycles cm^{-1}) and one very high (4.0 cycles cm^{-1}) spatial frequency Gabor. After 250 ms, one of the Gabors changed phase, shifting either to the left or to the right, and participants were required to make a speeded response regarding the direction. On one third of the trials, a modulating auditory sound matching the Gabor that changed was presented; on one third of the trials, a modulating auditory sound matching the other Gabor patch was presented; and on the remaining trials, no sound was presented. They reported that participants responded faster on trials on which the modulating auditory sound matched the Gabor patch that changed than when the modulating auditory sound matched the Gabor patch that did not. Since participants were instructed to ignore the sounds, Guzman-Martinez et al. concluded that attention was automatically captured by the modulating auditory sound to the matched visual spatial frequency. However, one may argue that the effect may have been mediated by goal-directed behavior, since the auditory signal and Gabor patches both preceded the target phase-shift by 250 ms. This pretarget time window may have allowed attention to be voluntarily directed to the target location in time for the phase-shift and, consequently, improve detection (Posner, 1980). A further possibility is that both the Gabor patches and the auditory sounds were very obviously different, facilitating a simple categorical judgement (e.g., between “very low” and “very high”). Given this potential confound, here, we will use a series of visual search tasks to test the proposal that modulating auditory sounds automatically captures attention to matched spatial frequencies.

In addition to their audiovisual matching experiment, Guzman-Martinez et al. (2012) also reported matching between vision and touch, using an amplitude-modulated tactile vibration that varied in rate, analogous to an auditory modulation rate (see

Guzman-Martinez et al., 2012, Experiment 3). Although they reported this matching relationship, the authors did not investigate the ability of tactile cues to capture attention to matching visual frequencies. There is growing evidence for early sensory interactions between vision and touch, and it has been suggested that the reorganization of the visual cortex of blind individuals that occurs in the absence of visual input reinforces preexisting connections between the somatosensory and visual cortices (see a review by Pascual-Leone, Amedi, & Fregni, 2005). Consistent with this suggestion, studies have demonstrated that the visual cortex of blind individuals is activated during tasks such as Braille reading (Cohen et al., 1997; Sadato et al., 1996) and haptic object recognition (Amedi, Raz, Azulay, Malach, & Zohary, 2010). Moreover, the existence of networks between vision and touch in normally sighted individuals is supported by neuroimaging studies; visual areas have been shown to be activated during tactile orientation discrimination (Sathian & Zangaladze, 2002; Sathian, Zangaladze, Epstein, & Grafton, 1999), haptic object recognition (Amedi, Jacobson, Hendler, Malach, & Zohary, 2002; Amedi et al., 2010; Deibert, Kraut, Kremen, & Hart, 1999; James, Humphrey, Gati, Servos, Menon, & Goodale, 2002; Pietrini et al., 2004), and Braille reading in sighted subjects following 5 days of complete visual deprivation (Merabet et al., 2008). Furthermore, there is behavioral evidence for the influence of tactile inputs on visual perception. Exploration of a tactile grating has been shown to produce prolonged dominance and reduced suppression of a visual grating of matched orientation and spatial frequency during binocular rivalry (Lunghi & Alais, 2013; Lunghi, Binda, & Morrone, 2010). Given these strong links, we also test the effects of tactile cues on search for matched visual spatial frequencies.

In this study, we report matching relationships between auditory modulation rate and visual spatial frequency, as well as between tactile spatial frequency and visual spatial frequency. We then use these matches in a series of visual search tasks. We compare visual search performance in the presence of auditory, tactile, and bimodal (auditory and tactile) cues with performance in the absence of cuing. Our findings demonstrate that informative auditory, tactile, and bimodal cues can produce improved search performance, with tactile cuing being more effective than auditory cuing and with tactile performance consistent with efficient search for certain matched spatial frequencies. However, our results also clearly demonstrate that these cross-modal cuing effects result from top-down attentional control, rather than bottom-up attentional capture, since uninformative tactile cues have no influence on visual search.

Experiment 1

This experiment involved two separate matching tasks that were performed sequentially: the first to determine the subjective

matching relationship between tactile and visual spatial frequency and, following this, a second to replicate the previously reported matching relationship between visual spatial frequency and the rate of amplitude modulation of an auditory stimulus (Guzman-Martinez et al., 2012). Although Guzman-Martinez et al. reported matching between vision and touch, the stimulus was an amplitude-modulated tactile vibration that varied in rate, analogous to auditory amplitude modulation rate (see Guzman-Martinez et al., 2012, Experiment 3). In the experiments reported here, the tactile stimulus was a 3-D object; as a result, the frequency of this stimulus is defined spatially, a feature more directly linked to vision. In the first task, participants adjusted visual Gabors to match a range of tactile gratings that were presented. On each trial, they were instructed to explore a single tactile grating and were asked to adjust the spatial frequency of a Gabor patch presented on the screen until they felt the visual image matched the tactile grating they were touching. The tactile gratings used had spatial frequencies of 0.66, 1.00, 1.33, 2.00, 2.66, and 4.00 cycles cm^{-1} . In the second task, participants were presented with the visual Gabors that they had previously reported as matching the tactile gratings and varied the amplitude modulation rate of an auditory stimulus to match the visual stimulus. Consistent with Guzman-Martinez et al., we expected an approximately linear relationship in the visual–auditory matching task and also in the tactile–visual matching task.

Method

Participants

Eight individuals (4 female; mean age, 23 years; range, 18–37 years; 1 left-handed) participated; all reported normal hearing and normal or corrected-to-normal vision. Six participants were naïve as to the purpose of the experiment. The sample size was predetermined prior to data collection, on the basis of related work from our laboratory (Orchard-Mills, Van der Burg, & Alais, 2013).

Stimuli and apparatus

Visual stimuli were displayed on a 16-in. CRT monitor (100 Hz, 1,024 × 768) situated 50 cm from the participants in a dark sound-attenuated room. Experiments were conducted using an Apple MacPro running MATLAB (The MathWorks) with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). A box was placed on the table between the participant and the monitor. This box contained a stepper motor that rotated a circular platform on which the six tactile gratings were placed. The lid of the box contained a small aperture that allowed participants to explore a single grating at a time. Gratings were vertically oriented and located in line with the fixation cross in the middle of the display.

Participants used the index finger of their dominant hand to explore the grating and their nondominant hand to enter responses on a keyboard. The tactile stimuli were circular sinusoidal gratings produced using a 3-D printer, with spatial frequencies of 0.66, 1.00, 1.33, 2.00, 2.66, and 4.00 cycles cm^{-1} and a radius of 1.5 cm. The visual stimuli were high-contrast Gabor patches (0.73 contrast) with a standard deviation of 0.46° (i.e., 0.4 cm, for comparison with the tactile stimulus) surrounded by a gray ring (radius 1.1° , or 1 cm) drawn with a Gaussian profile and displayed with background luminance of 51.5 cd m^{-2} . The auditory stimulus was superthreshold (SPL 65 dB) white noise high-pass filtered at 3500 Hz. The noise varied in amplitude from 0 % to 100 % using a sine wave envelope, the frequency of which could be adjusted. It was presented through speakers located on either side of the monitor.

Design and procedure

For the tactile–visual matching task, each trial began with the presentation of a gray screen (51.5 cd m^{-2}), the motor rotated the platform on which the tactile gratings were mounted to present one of the gratings at the inspection location and participants were instructed to explore the tactile grating with the index finger of their dominant hand. A vertically oriented Gabor patch (randomly chosen spatial frequency of between 0.5 and 10 cycles cm^{-1}) was presented on the screen, and participants were asked to adjust the visual spatial frequency to match the tactile grating. Four designated keys allowed participants to increase or decrease the spatial frequency of the Gabor patch in large (1.0) or small (0.1) steps, with no limitations placed on the spatial frequency that could be selected. This task was unsped, and participants pressed the space bar when they were satisfied with the match made. Participants completed 6 practice trials to familiarize themselves with the task. This was followed by a single block in which each tactile grating was presented 10 times (60 trials in total).

For the visual–auditory matching task, each trial began with the presentation of a gray screen, after which a vertically oriented Gabor patch, one of the matches reported by the participant in the tactile–visual matching experiment, appeared. A modulating auditory sound was presented, and participants were asked to adjust the rate of the modulation to match the Gabor patch. The starting modulation rate was randomized for each trial. Participants could increase or decrease the rate of the modulation in 0.5 Hz steps by pressing the up and down arrow keys, respectively. A lower limit of 0.5 Hz was placed on the modulation rate, however no upper limit was used. This task was unsped, and participants pressed the enter key when they were satisfied with the match made. Participants completed 6 practice trials to familiarize themselves with the task. This was followed by a single block

in which each Gabor was presented 10 times in random order (60 trials in total).

Results and discussion

The mean match was taken for each individual’s responses for each spatial frequency; these are shown in Fig. 1. The line of the best fit to the mean data is also shown for each of the matching relationships, generated using linear regression. The visual–tactile and visual–auditory matching tasks both demonstrate a linear relationship (see also Guzman-Martinez et al., 2012).

Each individual’s matching data were fitted with a linear function to produce individual matching relationships. These functions were subsequently used to assess the efficacy of cues to the matched spatial frequency in the visual search tasks in the following experiments.

Experiment 2

In this experiment, the visual search paradigm was used to investigate whether search is influenced by the matching relationships between visual and tactile spatial frequency and between auditory amplitude modulation rate and visual spatial frequency. In a cluttered visual display, participants searched for a ring with a horizontal notch (the target ring) among many rings with obliquely oriented notches (distractor rings, ±15° from horizontal; see Fig. 2 for an example display). Notches could be located on either the left or the right side of the ring, and participants were required to indicate, as quickly and accurately as possible, the location of the notch on the target ring (i.e., left or right side). Inside each ring was a Gabor patch, and the number of items in the display (either six or nine) was

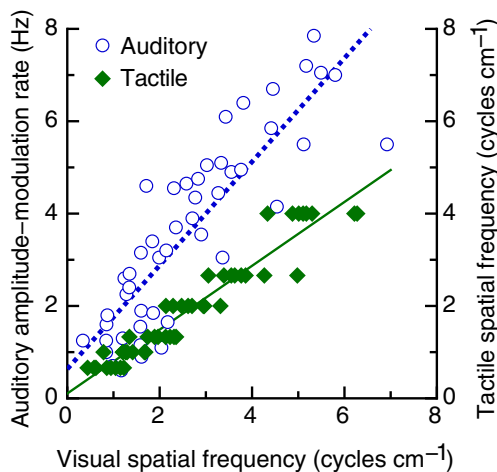


Fig 1 Frequency matching data from Experiment 1, showing individual mean responses and lines of best fit for visual–auditory ($r^2 = .74$) and visual–tactile ($r^2 = .92$) matching

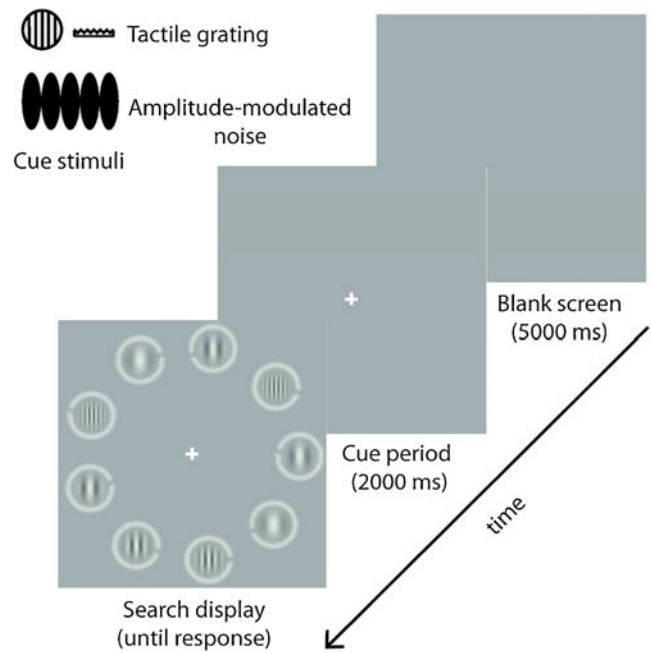


Fig. 2 Schematic representation of the time sequence of a trial in Experiments 2, 3, and 5. In the cued conditions, participants were given a 2,000-ms cue period before the search display appeared. Participants searched for the target ring, which had a horizontally aligned notch, among distractor rings with notches ±15° from horizontal. Participants were required to respond to the location of the notch (either left or right). Each ring contained a Gabor patch. In this set size of nine, three spatial frequencies in the distractor rings were repeated; in a set size of six, all Gabors were unique. The display example is drawn to scale; however, rings are thicker and brighter than the ones used in the task for illustrative purposes. There were four conditions used: no cue, auditory, tactile, and bimodal (both auditory and tactile). When cues were presented, they always matched the spatial frequency of the Gabor patch in the target ring. Participants were instructed that the cues were always informative and were told to use this information to speed their search

manipulated to assess search efficiency. Participants performed this task under four conditions: (1) with auditory cues, (2) with tactile cues, (3) with bimodal (auditory and tactile) cues, and (4) without any cues. When cues were present, they always matched the spatial frequency of the Gabor inside the target ring, and each individual’s own tactile–visual and visual–auditory matching relationships from Experiment 1 were used. Critically, participants could perform the search task without cues to the Gabor within the target ring (see also a similar methodology used in Theeuwes & Van der Burg, 2007, Theeuwes & van der Burg, 2008, Theeuwes & Van der Burg, 2011). Participants were informed that these cues were always valid and were told to use the information to facilitate their search for the target ring. Participants were given 2,000 ms to process the cue prior to the appearance of the display and were allowed to explore the tactile grating during search. The auditory signal was presented until participants made a response. If the cue can guide attention toward the cued spatial frequency and, therefore, the target ring, we expect more efficient visual

search (shallower search slopes) for the cued conditions, as compared with the no-cue condition.

Method

Participants

The same individuals as those in Experiment 1 participated.

Stimuli and apparatus

The experimental setup was the same as in Experiment 1. The tactile gratings had spatial frequencies of 0.66, 1.00, 1.33, 2.00, 2.66, or 4.00 cycles cm^{-1} . For the visual spatial frequency of each Gabor patch, we used the individual's matches from Experiment 1. These were derived from the best-fitting linear function to their responses for the visual spatial frequency. For the amplitude modulation rate of the auditory stimulus, we again used the individual's matches from Experiment 1. These were derived from the best-fitting linear function to their responses for the auditory modulation rate.

Design and procedure

Each trial began with a 5,000-ms blank gray screen; during this time, the stepper motor rotated the platform on which the tactile gratings were mounted to present a single grating in the tactile cue condition. A fixation cross then appeared, signaling the beginning of the 2,000-ms cue period. This cue period was chosen because it allowed sufficient time for participants to register the signal, move their finger to touch the grating, and adequately explore the grating to process the spatial frequency presented. In the tactile and bimodal cue conditions, participants were instructed to explore the tactile grating with their dominant hand until they had located the target ring. In the auditory and bimodal cue conditions, the auditory stimulus began upon presentation of the fixation cross and was presented until participants made their response. In all cue conditions, the cue matched the visual spatial frequency of the Gabor patch in the target ring with 100 % validity. After the cue period, the visual search display appeared, with either six or nine rings each containing a vertically oriented Gabor patch (see Fig. 2). The rings and Gabor patches were equally spaced around fixation on an imaginary circle (radius 6°). Each gray ring was drawn with a Gaussian profile (radius 1.1°) and had a small notch removed that matched the background color. The target ring was always present and had a notch that was horizontally aligned, whereas the notches in the distractor rings deviated $\pm 15^\circ$ from horizontal. Each notch was randomly placed on either the left or the right, and participants were asked to respond, as quickly and accurately as possible, as to the location of the notch on the target ring (left or right side) by pressing the “f” or “h” key, respectively, using their nondominant hand. Each

ring contained a Gabor patch (0.73 contrast, standard deviation of 0.4 cm or 0.46 c/d). In the set size of six, all the Gabors had unique spatial frequencies. In the set size of nine, three of the spatial frequencies in the distractor rings (randomly chosen) were each replicated once, ensuring that the differences in spatial frequency between the Gabors in the display were constant across both set sizes. The spatial frequency of the Gabor in the target ring was always unique and matched the cue (if present). However, importantly, the search task was possible without the cue. The notch location for the target ring, the visual spatial frequency within the target ring and set size were randomly chosen for each trial and balanced within blocks, with the limitation that the spatial frequency inside the target ring was never repeated in the next trial, to avoid any intertrial priming (Theeuwes & Van der Burg, 2011). An example of the visual display for a set size of nine is shown in Fig. 2.

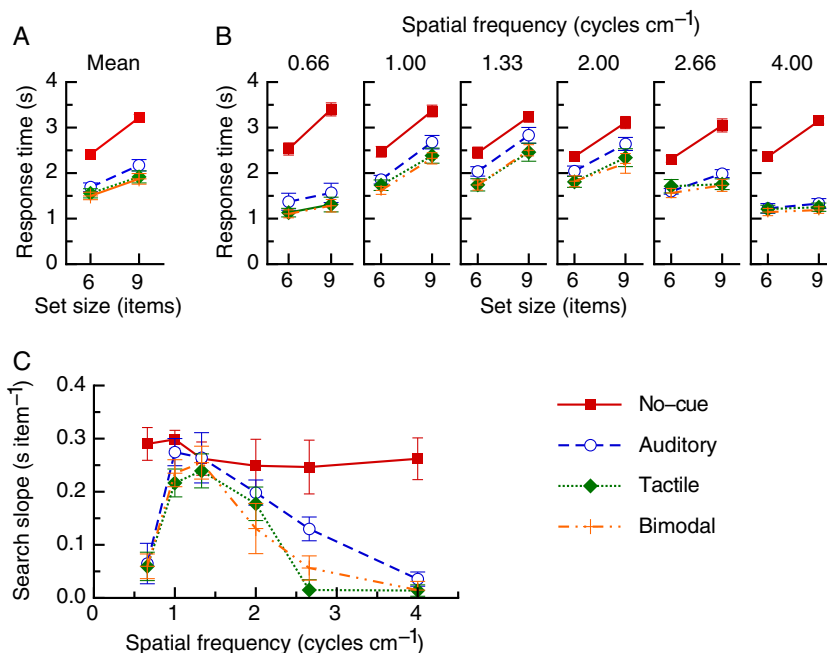
Each cue condition was presented in separate blocks. Each participant completed four sessions, each containing one block of each condition (total, 1,536 trials). The order of cue conditions was counterbalanced within sessions and between participants. Prior to the first session, a practice block of 15 trials for each cue condition was completed. Breaks were allowed as required between each block. Participants were instructed that they would be presented with their perceptual matches from Experiment 1 and that the cues would always be valid and were told to use this information to speed their search.

Results and discussion

The overall error rate was low (1.9 %); therefore, no further analysis of errors was conducted, and all erroneous trials were excluded from the analysis. Response times (RTs) greater than 10 s and those that differed more than 2.5 standard deviations from the mean for each participant for each condition were also excluded from the analysis (1.7 % of responses). Since RTs may have been influenced by the cue providing a temporal marker (see Los & Van der Burg, 2013) or improving alertness, search slopes were calculated for each individual for each spatial frequency for each condition. Figure 3a shows overall correct mean RTs as a function of set size and cue condition.

Search slopes were subjected to an ANOVA with spatial frequency and cue condition as within-subjects variables; the Huynh–Feldt correction was used when appropriate. The ANOVA yielded a significant effect for the cue condition, $F(3, 21) = 43.0$, $p < .001$. The effect of the cue condition varied across the spatial frequencies tested. This was confirmed by a significant two-way interaction between cue condition and spatial frequency, $F(15, 105) = 3.3$, $p < .001$. The change in RTs and search slopes over the range of spatial frequencies tested is shown in Fig. 3b, c. Pairwise comparisons, Bonferroni corrected, for each cue condition showed significant differences between the auditory cue and no-cue

Fig. 3 Data from Experiment 2. **a** Overall correct mean response times (RTs) for the no-cue, auditory, tactile, and bimodal cue conditions for each set size. **b** Correct mean RTs for each target spatial frequency for each set size and cue condition. **c** Mean search slopes for each target spatial frequency for each cue condition. Error bars represent ± 1 SEM



conditions, $p = .006$, between the tactile cue and no-cue conditions, $p = .001$, and between the bimodal cue and no-cue conditions, $p = .002$. This demonstrates that search was more efficient in the presence of cues (auditory, 161 ms per item; tactile, 120 ms per item; bimodal, 125 ms per item) than in the absence of cues (268 ms per item). Interestingly, the auditory and tactile cue conditions were significantly different, $p = .001$, and there was a significant difference between the auditory and bimodal conditions, $p = .002$, but not between the tactile and bimodal cues, $p = .99$. This demonstrates that there is no additional gain for combining auditory and tactile cues, since bimodal cuing performance does not exceed that of tactile cuing performance. The absence of improvement for bimodal cuing in visual search and cross-modal attentional capture has been reported previously (Ngo & Spence, 2010; Santangelo, Van der Lubbe, Belardinelli, & Postma, 2006).

Pairwise comparisons for each spatial frequency, Bonferroni corrected, showed reliable differences between the auditory cue condition and no-cue condition for the lowest and highest spatial frequencies of 0.66 and 4.0 cycles cm⁻¹ ($p = .036$ and $p = .006$, respectively), but not for the intermediate spatial frequencies of 1.0, 1.33, 2.0, and 2.66 cycles cm⁻¹ (all p values > .45): for these four spatial frequencies, auditory cues were no more effective than the no-cue condition. For the tactile cue condition, the difference was significant for the lowest and two highest spatial frequencies of 0.66, 2.66, and 4.0 cycles cm⁻¹ (all p values < .02), but not for the intermediate spatial frequencies of 1.00, 1.33, and 2.0 cycles cm⁻¹ (all p values > .07). For the bimodal cue condition, the difference was significant for the lowest and highest spatial frequencies of 0.66 and 4.0 cycles cm⁻¹ (all p values < .007).

Interestingly, in the tactile condition, the search slopes were not significantly greater than zero for target spatial frequencies of 2.66 (15 ms per item), $t(7) = 0.8$, $p = .4$, and 4 (14 ms per item), $t(7) = 1.4$, $p = .2$, cycles cm⁻¹. In the bimodal condition, the search slope was not significantly greater than zero for a target spatial frequency of 4 cycles cm⁻¹ (15 ms per item), $t(7) = 1.0$, $p = .4$. In the auditory condition, the search slope was not significantly greater than zero for a target spatial frequency of 0.66 cycles cm⁻¹ (65 ms per item), $t(7) = 1.8$, $p = .1$. In contrast, all other search slopes were significantly greater than zero (all p values < .05).

The motivation for this experiment was to investigate whether matching relationships reported by Guzman-Martinez et al. (2012) would influence visual search. We used informative and valid cues to ensure participants attended to the information, giving maximum chance for the cues to improve search. Overall, these results demonstrate that informative auditory cues can improve visual search efficiency to matched spatial frequencies only in a limited manner, with only the lowest and highest spatial frequencies benefitting from an auditory cue. The four intermediate spatial frequencies did not show significant search benefits from auditory cuing. This indicates that although participants were motivated to use the auditory cues, they were unable to use them effectively for most of the spatial frequencies presented. This does not support the proposal by Guzman-Martinez et al. that modulating sounds produce attentional capture to matched spatial frequencies.

Tactile cues to matched visual spatial frequencies also influenced visual search, with effects greater in strength and specificity than auditory cues. Tactile cues produced search benefits for three out of six of the spatial frequencies in the

display. Interestingly, two of the search slopes for the tactile condition could be considered pop-out search, since search slopes were not significantly different from zero. This shows that tactile cuing can produce efficient search and, therefore, it may be that, for some spatial frequencies, tactile cues do produce attentional effects. If the cues from audition and touch were to summate, we would expect search performance to be best for bimodal cues; however, the results demonstrate that bimodal cuing is no better than tactile cuing. This also suggests that tactile cues are more effective than auditory cues.

Experiment 3

In Experiment 2, search was improved for the lower and higher ends of the range of spatial frequencies, but not for the intermediate frequencies. This raises the question of whether the failure to find search performance improvements for the intermediate spatial frequencies was due to the spatial frequency differences between the Gabor patches in the display. It is possible that the informative tactile cues were, in fact, very effective (as is the case for the highest spatial frequencies), but due to the relatively small differences between the target and distractor Gabors, participants were unable to directly guide attention to the target spatial frequency for the intermediate frequencies. This experiment was designed to investigate whether an informative and valid visual cue would produce search benefits for all spatial frequencies in the display, not just at the upper and lower limits. In this experiment, prior to the search display, participants were presented with the Gabor that would appear in the target ring as a cue. Previous studies have shown that this improved visual selection in a visual search task when using a color cue (Theeuwes & Van der Burg, 2007, 2011). Participants completed the tactile cue condition from Experiment 2 as a comparison with the visual cue condition. As in Experiment 2, participants were instructed to use the information, since the cue always matched the Gabor in the target ring. If the visual cue produces improved search for all spatial frequencies, we can conclude that it is not visual spatial resolution that limited tactile performance at intermediate frequencies in Experiment 2 but that the tactile cues did not effectively guide search to the matched spatial frequency.

Method

The experiment was identical to Experiment 2, using the same participants, except for the following changes. Only two cue conditions were used, a tactile cue condition and visual cue condition. In the tactile condition, each trial was identical to the tactile condition in Experiment 2. In the visual condition, the Gabor that would be presented inside the target ring was used as the cue stimulus. It was displayed centrally for

1,000 ms during the cue period, followed by a fixation cross for 1,000 ms. The visual cue was removed from the screen to avoid any apparent motion from the cue to the Gabor patch within the target ring (see Theeuwes & Van der Burg, 2007). Participants were informed that they would be presented with 100 % valid cues and were told to use this to speed their search. Each participant completed four blocks of 96 trials for each condition (total, 768 trials) presented in counterbalanced order.

Results and discussion

The overall error rate was low (1.1 %); therefore, no further analysis of errors was conducted, and all erroneous trials were excluded from the analysis. RTs greater than 10 s and those that differed more than 2.5 standard deviations from the mean for each participant for each condition were also excluded from the analysis (1 % of the trials). The overall correct mean RTs are shown in Fig. 4a, correct mean RTs for each spatial frequency in Fig. 4b, and mean search slopes for each spatial frequency in Fig. 4c.

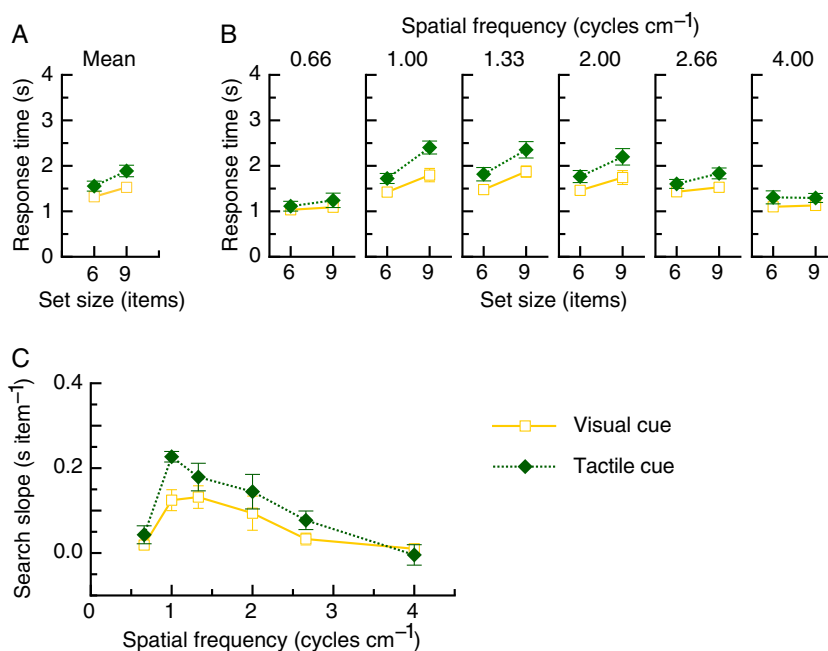
Search slopes were subjected to an ANOVA with spatial frequency and cue condition as within-subjects variables. This yielded a significant effect for the cue condition, $F(1, 7) = 6.7$, $p < .05$, since search slopes were lower for visual cues (69 ms per item) than for tactile cues (111 ms per item). The interaction between spatial frequency and cue was not significant, $F(5, 35) = 2.4$, $p = .07$.

Search was very efficient for the lowest and highest spatial frequencies in both the tactile cue (43 ms per item and -4 ms per item, respectively) and visual cue (19 ms per item and 10 ms per item, respectively) conditions. Although it is quite likely that a coarser sampling of the spatial frequencies presented in the display would produce efficient search for all items in the display, visual cuing was more efficient than tactile cuing for the intermediate spatial frequencies in the display. This indicates that search performance for tactile cuing was not limited by the visual display; search can be more efficient if sufficient information is provided, as was seen with visual cues. Previous studies using informative cues have demonstrated that performance in conjunction search using word cues did not reach the performance produced with cuing using the visual target itself (Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004). The findings reported here demonstrate that although tactile cues provide useful top-down information for participants to search for the matched spatial frequency, participants are able to use visual information more effectively.

Experiment 4

In Experiments 2 and 3, the tactile cues to matched visual spatial frequencies were informative and 100 % valid. This

Fig. 4 Data from Experiment 3. **a** Overall correct mean response times (RTs) for the visual and tactile conditions for each set size. **b** Correct mean RTs for each target spatial frequency for each set size and cue condition. **c** Mean search slopes for each target spatial frequency for each cue condition. Error bars represent ± 1 SEM



demonstrates that participants could use this information to speed their search and that search was highly efficient for the two highest spatial frequencies. However it is unclear from this experiment whether search benefits were due to top-down attention or bottom-up capture, as was suggested by Guzman-Martinez et al. (2012). In this experiment, we investigate whether uninformative tactile information influences search for a matched visual frequency.

The participants' task was to search for the target ring and to respond to the location of the notch (as in Experiment 2). However, this experiment used only two cue conditions (a no-cue condition and a tactile condition). In the tactile condition, the tactile cue always had the same spatial frequency (4.00 cycles cm⁻¹). We decided to use this spatial frequency, since this grating produced the most efficient visual search performance in Experiment 2. In contrast to Experiment 2, where the tactile cues always matched the Gabor in the target ring, in this experiment the tactile cue matched the Gabor in the target ring in only one out of six trials, corresponding to chance level. To ensure that participants explore the grating on each trial, it was obliquely oriented in 25 % of trials and vertical in the remaining trials. Participants made a speeded response to the target ring and, following this, an unspeeded response to the tactile orientation. If exploring the tactile grating automatically captures attention to the matched visual Gabor patch, search for the target ring containing this spatial frequency will be faster, as compared with the unmatched spatial frequencies. In contrast, if no difference in search times were found, it would suggest that voluntary attention is required to produce the search benefits demonstrated in Experiment 2. Prior to the search task, participants performed the same matching task as in Experiment 1, using only the 4.00 cycles cm⁻¹ grating, this

ensured that the display contained a Gabor patch that they reported as matching the tactile grating (as in Experiment 2).

Method

The experiment was identical to Experiment 2, except for the following changes.

Participants

Eight new individuals (3 female; mean age, 23 years; range, 18–33 years; 2 left-handed) participated; all participants were naïve as to the purpose of the experiment.

Design and procedure

Participants performed a matching task prior to the search task. The matching task was identical to that in Experiment 1; however, only the 4.00 cycles cm⁻¹ grating was presented. The mean response of 10 trials was taken as the participant's perceptual match.

In the search task, the search display was identical to the display in Experiment 2, except that the set size was always six. For the six Gabors, one was the individual participant's match for the grating presented, which was determined in the matching task of this experiment; the other five were the overall mean responses for the gratings of 0.66, 1.00, 1.33, 2.00, and 2.66 cycles cm⁻¹ from Experiment 1, which were 0.9, 1.4, 2.1, 2.8, and 3.8 cycles cm⁻¹. The search task was identical to the task in Experiment 2. Participants searched for the target ring and responded as quickly and accurately as to the location of the notch.

In the tactile condition, the tactile grating was always $4.00 \text{ cycles cm}^{-1}$. It therefore matched the Gabor in the target ring in only one out of six trials (16.6 % validity). In five out of six trials, it matched a Gabor in a distractor ring. Therefore, the tactile cue was uninformative, since it matched the Gabor in the target ring at chance level. To ensure that participants processed the tactile cue, on each trial participants also performed a tactile orientation task. The grating was obliquely oriented (± 5.4 , ± 16.2 , or $\pm 45^\circ$ from vertical) on 25 % of trials and vertical on the remaining trials. After making a speeded response to the target ring, participants then made an unspeeded response indicating whether the grating was vertically oriented or otherwise, using the “f” or “h,” keys, respectively. This ensured that participants processed the tactile information on each trial. Each participant completed one block of 144 trials for each cue condition (total, 288 trials). The order of presentation of the blocks was counterbalanced between participants. Each condition began with a practice block of 15 trials, and breaks were allowed every 36 trials.

Results and discussion

The mean matching responses are shown in Fig. 5b, along with the mean responses for the same spatial frequency grating from Experiment 1.

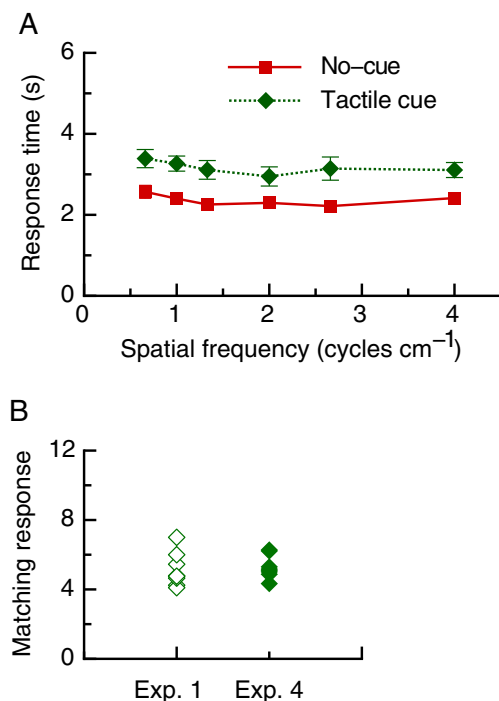


Fig. 5 Data from Experiment 4. **a** Overall correct mean response times for the no-cue and tactile conditions for each target spatial frequency. **b** Data from Experiments 1 and 4, showing individual mean responses for the visual spatial frequency match for a tactile grating of $4.0 \text{ cycles cm}^{-1}$. Error bars represent $\pm 1 \text{ SEM}$

There was no significant difference when the matching responses were compared using a *t*-test, $t(7) = 0.3$, $p = .8$. This shows consistency in the matching relationship, even when participants match only one tactile grating to visual spatial frequency. The performance on the grating orientation task exceeded 90 % at an orientation of $\pm 45^\circ$, demonstrating that participants actively engaged with the tactile cues on each trial.

The overall error rate for the search task was low (1.2 %), and therefore, no further analysis of errors was conducted. All erroneous trials and trials on which the grating was obliquely oriented were excluded from the analysis. RTs greater than 10 s and those that differed more than 2.5 standard deviations from the mean for each participant for each condition were also excluded from analysis (1.4 % of responses). Figure 5a shows the correct mean RTs as a function of spatial frequency. Response times were subjected to an ANOVA with spatial frequency and cue condition as within-subjects variables. This yielded a significant effect for the cue condition, $F(1, 7) = 20.3$, $p = .003$, since search times were shorter for the no-cue condition (2,358 ms) than for the tactile cue condition (3,160 ms). One feasible explanation for the search costs in the tactile condition, as compared with the no-cue condition, is that participants were still processing the tactile cue when the search display appeared. Although the main effect of spatial frequency was reliable, $F(5, 35) = 3.9$, $p = .008$, critically, no significant interaction between spatial frequency and cue was found, $F(5, 35) = 0.8$, $p = .5$. This indicates that search for the matched spatial frequency was no faster than that for the unmatched spatial frequencies.

This experiment demonstrates that when uninformative, tactile cues to matched visual spatial frequencies have no effect on visual search, and therefore, there is no evidence of attentional capture. These results suggest that voluntary goal-directed attention is required to produce the cross-modal search benefits seen in Experiment 2.

Experiment 5

In Experiment 5, we further examine whether the search benefits observed in Experiments 2 and 3 are due to a top-down effect or to a stimulus-driven effect (as proposed by Guzman-Martinez et al., 2012, for modulating sounds). There were two cue conditions used. In one condition, participants were presented with word cues giving the ordinal position of the spatial frequency of the Gabor in the target ring—for example, “One” if the Gabor had the lowest spatial frequency or “Six” if the Gabor had the highest spatial frequency. In the other cue condition, the letters “XXXX” were displayed, giving participants no information regarding the Gabor in the target ring. If the search benefits observed in Experiments 2 and 3 were due to top-down knowledge, we

would expect to observe similar search benefits for a word cue. In contrast, if the search benefits for cross-modal cues were due to a sensory effect, we would not expect a word cue to replicate the pattern of search benefits observed in Experiments 2 and 3.

Method

Participants

Eight new individuals (3 female; mean age, 22 years; range, 18–32 years) participated; all participants were naïve as to the purpose of the experiment.

Design and procedure

The experiment was identical to Experiment 2, except for the following changes. Only two cue conditions were used: a word-cue condition and a no-cue condition. In the word-cue condition, participants were presented with a written number representing the position of the Gabor in the target ring in the range of spatial frequencies in the display. If the target ring contained the Gabor with the lowest spatial frequency, the word “One” was displayed; if it was the second highest spatial frequency, the word “Two” was displayed; and so forth. In the no-cue condition, the letters “XXXX” were displayed to provide the same temporal cue as the word-cue condition. The cues were written in white, Helvetica font size 24. The cues were displayed centrally for 1,000 ms during the cue period, followed by a fixation cross for 1,000 ms. Participants were informed that in the word-cue condition, they would be presented with 100 % valid cues and were told to use this information to speed their search. Each participant completed 15 training trials, followed by a single blocks of 240 trials presented in random order.

Results and discussion

The overall error rate was low (5.0 %). RTs greater than 10 s and those that differed more than 2.5 standard deviations from the mean for each participant for each condition were also excluded from the analysis (1.8 % of the trials). The correct mean RTs as a function of the spatial frequencies tested are shown in Fig. 6.

RTs were subjected to an ANOVA with spatial frequency and cue condition as within-subjects variables. This yielded a significant effect for the cue condition, $F(1, 7) = 33.3, p = .01$, since RTs were shorter for the word-cue condition (2,144 ms) than for the no-cue condition (2,900 ms). The interaction between spatial frequency and cue was significant, $F(5, 35) = 5.6, p = .01$. This indicates that the effect of the cue condition was dependent on the spatial frequency. This was confirmed by pairwise comparisons for each spatial

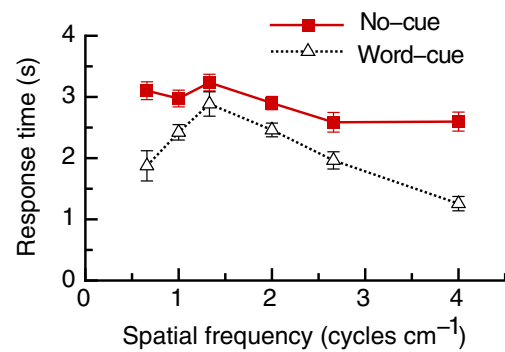


Fig. 6 Data from Experiment 5, showing correct mean response times for each target spatial frequency for each cue condition. Error bars represent ± 1 SEM

frequency, Bonferroni corrected, demonstrating reliable differences between the word-cue condition and no-cue condition for the spatial frequencies of 0.66, 1.00, 2.00, 2.66, and 4.0 cycles cm^{-1} (all p values $< .02$), but not 1.33 cycles cm^{-1} ($p = .25$).

There were more errors and exclusions in the no-cue condition (4.2 %) than in the word-cue condition (2.6 %); however, when these were subjected to the same ANOVA, no significant effects were found (all p values $> .15$). This indicates that a speed–accuracy trade-off does not explain the RT benefits reported for the word-cue condition.

These results show a similar pattern of RT benefits to those seen with both the cross-modal (Experiment 2) and visual (Experiment 3) cue conditions, with the greatest search benefits seen for the lowest and highest spatial frequency items in the display. This suggests that the search benefits reported are not stimulus driven but, rather, due to top-down knowledge.

General discussion

In the present study, Experiment 1 replicated previous findings that participants match the visual spatial frequency of a Gabor patch with auditory amplitude modulation rate with a linear relationship (Guzman-Martinez et al., 2012). It also showed that participants match the frequency of a tactile grating with visual spatial frequency of a Gabor patch with a linear relationship. Recently Guzman-Martinez et al. suggested that the audiovisual matching relationship resulted from repeated concurrent multisensory inputs during manual exploration of surfaces. They reported findings that suggested that this matching relationship produced automatic attentional effects. Using a series of visual search tasks, we investigated whether automatic attentional effects occurred between visual spatial frequency and auditory amplitude modulation rate or tactile spatial frequency.

Experiment 2 demonstrated that informative and valid auditory cues produce search benefits. However, search was

improved only for the lowest and highest spatial frequencies in the display. We suggest that the modulating auditory sounds provide only a broad relative cue, with slow modulating sounds cuing coarse spatial frequencies and faster modulating sounds cuing fine spatial frequencies, rather than specific matches. This would account for both the search benefits for low and high spatial frequencies and the failure to find benefits for intermediate spatial frequencies. This suggests that auditory cues facilitate search performance through top-down attentional control rather than automatic attention capture, as was suggested by Guzman-Martinez et al. (2012).

The results are in line with our previous study (Orchard-Mills et al., 2013). This study investigated the proposed attentional effects of auditory cues to matched spatial frequencies. The visual search display in that study contained much finer sampling of spatial frequency, and participants were presented with nine different matches, as compared with the six used in the present study. However, despite these differences, the results were very similar. Although informative auditory cues did produce search benefits, these were absent for intermediate spatial frequencies in the display, and when uninformative, auditory cues had no effect on visual search. Furthermore, auditory cues that violated the matching relationship but were still informative produced the same pattern of search benefits as that seen for matched cues. Taken together, these findings indicate that auditory cues produce search benefits through top-down attention, not bottom-up capture.

The present study extends our previous study by showing that informative and valid tactile cues produced search benefits (Experiments 2 and 3). Interestingly, tactile cues produced greater search benefits, as compared with auditory cues, since overall search slopes were shallower for tactile cues than for auditory cues. Tactile cues produced significant benefits, as compared with the no-cue condition, for three out of the six spatial frequencies tested. Furthermore, tactile cues produced highly efficient search slopes for the two highest spatial frequencies; both were not significantly different from zero (15 and 14 ms per item, respectively). However, Experiment 4 showed that when uninformative, the tactile cue had no influence on search times. This suggests that although matching between visual and tactile spatial frequency can produce search benefits, the effects were not due to automatic capture (see Theeuwes, 2010, for a review regarding attentional capture) but, instead, were driven by top-down knowledge.

The finding that the greatest search benefits were seen for the lowest and highest spatial frequencies is consistent with visual search literature regarding target–distractor relationships (Becker, 2010; Wolfe & Horowitz, 2004). Becker suggested a relational selection mechanism for visual search, which predicts that items that are found at the end of a relational vector are found more efficiently. This incorporates previous feature-based accounts of linearly separability; for example, in a display defined by color, search has been shown

to be efficient if all the elements extend along a line in two-dimensional color space and the target is located at one end (Bauer, Jolicoeur, & Cowan, 1996; D’Zmura, 1991). Similarly for orientation, search for the steepest item has been shown to be more efficient than that for the third steepest item (Wolfe, Friedman-Hill, Stewart, & O’Connell, 1992). Since auditory cues were effective only for the lowest and highest spatial frequencies, one could explain the search benefits using a top-down strategy where participants search for the lowest spatial frequency items for a slowly modulating sound and the opposite for a fast modulating sound. Since these are the most efficiently found visually, this strategy is effective.

It is possible that this broad, relative strategy is also used for tactile cues. However, since search is generally inefficient if, along the relevant feature dimension, the distractors have values that flank that of the target (Wolfe & Horowitz, 2004), the finding that tactile cues produced efficient search for a spatial frequency that is not linearly separable from other items is inconsistent with this account. Instead, it implies that participants were searching for the specific target. This suggests that tactile cues provide more frequency-specific information than does audition, which may result from the acute spatial resolution of touch (Loomis, 1979) and the fact that tactile and visual modalities are both spatially defined senses that can be directly mapped together. Since similar patterns of search benefits were seen for cross-modal (Experiment 2), visual (Experiment 3), and word cues (Experiment 5), it is likely that the cross-modal cues facilitated search through top-down knowledge (e.g., due to visual imagery or through a semantic/verbal cue to “low” or “high” items in the display). Although tactile cuing produced efficient search for some spatial frequencies, the results of Experiment 3 demonstrated that visual cues produced more efficient search slopes than did tactile cues. This may have occurred because the visual cue provided a more accurate or more detailed representation of the target (Wolfe et al., 2004) or may have been a result of visual priming (Theeuwes, Reimann, & Mortier, 2006; Theeuwes & Van der Burg, 2007; 2011).

Interestingly, in Experiment 2, the bimodal cuing performance was greater than that of auditory cuing but no better than that of tactile cuing. Since Experiment 3 demonstrated that visual cuing produced greater search benefits than did tactile cuing, the failure to find larger search benefits for bimodal cues suggests that bimodal cues do not provide any further information than do tactile cues alone. One possible explanation for the absence of a stronger bimodal cuing effect is that the cues are presented in different frames of reference; the tactile cues were spatial and the auditory cues temporal, and this may have prevented summation. However, we suggest that this occurs because auditory cues provide only broad relative information, allowing participants to search for lower or higher spatial frequency targets, rather than cuing a specific spatial frequency match.

The fact that we did not find any evidence for automatic capture with tactile cues contrasts with previous studies that have suggested that touch influences visual processing in early sensory areas (Lunghi & Alais, 2013; Lunghi et al., 2010). Lunghi and Alais showed that exploration of a tactile grating produced prolonged dominance and reduced suppression of a grating of matching orientation during binocular rivalry. This effect was tuned for both orientation and spatial frequency (Lunghi & Alais, 2013), suggesting that it occurred through interactions in early visual areas that are selective for these features. This is in contrast to the dependence on voluntary goal-directed behavior found in this report, which suggests higher-level processes.

One key difference between the tactile experiments reported here and that of Lunghi et al. (2010) is the spatial proximity between the visual and tactile stimuli used. In the present experiment, the tactile cues were presented on a table in front of and below the visual display. In contrast, in the experiment reported by Lunghi et al., the tactile gratings were spatially collocated with the visual stimuli using mirrors. Spatial proximity is a critical factor for many multisensory interactions (Alais, Newell, & Mamassian, 2010; Stein & Meredith, 1993), although there are also many examples of cross-modal interactions for which spatial proximity is not a prerequisite (Murray et al., 2005; Olivers & Van der Burg, 2008; Shams, Kamitani, & Shimojo, 2000; Van der Burg, Cass, Olivers, Theeuwes, & Alais, 2008a, 2008b; Violentyev, Shimojo, & Shams, 2005).

Another key difference between our experiments and that of Lunghi et al. (2010; Lunghi & Alais, 2013) is the level of complexity in the display. Lunghi et al. and Lunghi and Alais presented only one visual item, whereas the experiments reported here used an array of visual items. In our experiments, attention was presumably diffusely directed over the whole display before either the target was correctly identified (for those spatial frequencies that produced efficient search) or attention was sequentially directed to one of the elements at a time (in the case of inefficient search). In the experiments reported by Lunghi et al. and Lunghi and Alais, attention was directed to a single location, where both the visual and tactile inputs were collocated; consequently, attention was always directed at spatially matched inputs. This may be critical to the low-level interactions between the two modalities reported by these studies. Many studies have reported that attention is a prerequisite for observing multisensory integration (Alsuis, Navarra, Campbell, & Soto-Faraco, 2005; Alsuis, Navarra, & Soto-Faraco, 2007; Fujisaki, Koene, Arnold, Johnston, & Nishida, 2006; Talsma, Doty, & Woldorff, 2007), although others report findings to the contrary (Van der Burg et al., 2008a; and see Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010, for a recent review regarding the role of attention in multisensory integration).

In this article, we have reported two matching relationships: a matching relationship between tactile and visual

spatial frequency and one between auditory amplitude modulation rate and visual spatial frequency. These matching relationships were recently proposed to result from recurring concurrent multisensory inputs during manual exploration of surfaces (Guzman-Martinez et al., 2012). Guzman-Martinez et al. reported findings that suggested the matching relationship between audition and vision produced automatic attentional capture (although see alternative explanations in Orchard-Mills et al., 2013). Using a series of visual search tasks, our results have shown that informative and valid auditory, tactile, and bimodal cues can produce more efficient visual search for matched visual spatial frequencies. However, auditory cues improved search performance only for the lowest and highest spatial frequencies. Tactile cues were shown to produce search benefits greater in both strength and specificity, with highly efficient search for some spatial frequencies. Most important, however, we show that when uninformative, a tactile cue produces no effect on visual search for a matched spatial frequency. Furthermore, using the same display, tactile cues were found to be less effective than visual cues, and word cues were found to produce a similar pattern of search benefits. These findings suggest that the improvements in search performance produced by both auditory and tactile cues are achieved through voluntary goal-directed behavior, rather than automatic attentional capture.

References

- Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory processing in review: from physiology to behaviour. *Seeing and perceiving*, 23(1), 3–38. doi:10.1163/187847510X488603
- Alsuis, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research*, 183(3), 399–404. doi:10.1007/s00221-007-1110-1
- Alsuis, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audio-visual integration of speech falters under high attention demands. *Current Biology*, 15(9), 839–843. doi:10.1016/j.cub.2005.03.046
- 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(11), 1202–1212.
- Amedi, A., Raz, N., Azulay, H., Malach, R., & Zohary, E. (2010). Cortical activity during tactile exploration of objects in blind and sighted humans. *Restorative Neurology and Neuroscience*, 28(2), 143–156. doi:10.3233/RNN-2010-0497
- Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly separable from distractors. *Vision research*, 36(10), 1439–1466. doi:10.1016/0042-6989(95)00207-3
- Becker, S. I. (2010). The role of target-distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology: General*, 139(2), 247–265. doi:10.1037/a0018808
- Bolia, R. S., D'Angelo, W. R., & McKinley, R. L. (1999). Aurally aided visual search in three-dimensional space. *Human Factors*, 41(4), 664–669.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. doi:10.1163/156856897X00357
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica United with Acustica*, 86(1), 117–128.

- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, 25(5), 975–979. doi:10.1121/1.1907229
- Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Falz, L., Dambrosia, J., ... Hallett, M. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature*, 389(6647), 180–183. doi:10.1038/38278
- Deibert, E., Kraut, M., Kremen, S., & Hart, J. (1999). Neural pathways in tactile object recognition. *Neurology*, 52(7), 1413–1417.
- D'Zmura, M. (1991). Color in visual search. *Vision research*, 31(6), 951–966.
- Fujisaki, W., Koene, A., Arnold, D., Johnston, A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. *Proc Biol Sci*, 273(1588), 865–874. doi:10.1098/rspb.2005.3327
- 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 : CB*, 22(5), 383–388. doi:10.1016/j.cub.2012.01.004
- 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(2), 252–259. doi:10.1016/j.actpsy.2010.07.017
- 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(7), 1736–1741. doi:10.3758/APP.72.7.1736
- Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, 15(3), 548–554.
- 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(10), 1706–1714. doi:10.1016/S0028-3932(02)00017-9
- Jones, C. M., Gray, R., Spence, C., & Tan, H. Z. (2008). Directing visual attention with spatially informative and spatially noninformative tactile cues. *Experimental Brain Research*, 186(4), 659–669. doi:10.1007/s00221-008-1277-0
- Loomis, J. M. (1979). An investigation of tactile hyperacuity. *Sensory Processes*, 3(4), 289–302.
- Los, S. A., & Van der Burg, E. (2013). Sound speeds vision through preparation, not integration. *Journal of Experimental Psychology: Human Perception and Performance*. doi:10.1037/a0032183
- Lunghi, C., Binda, P., & Morrone, M. C. (2010). Touch disambiguates rivalrous perception at early stages of visual analysis. *Current Biology : CB*, 20(4), R143–R144. doi:10.1016/j.cub.2009.12.015
- Lunghi, C., & Alais, D. (2013). Touch interacts with vision during binocular rivalry with a tight orientation tuning. *PLoS ONE*, 8(3), e58754.
- McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. R. (2004). Visual skills in airport-security screening. *Psychological Science*, 15(5), 302–306. doi:10.1111/j.0956-7976.2004.00673.x
- Merabet, L. B., Hamilton, R., Schlaug, G., Swisher, J. D., Kiriakopoulos, E. T., Pitskel, N. B., ... Pascual-Leone, A. (2008). Rapid and reversible recruitment of early visual cortex for touch. *PLoS One*, 3(8), e3046. doi:10.1371/journal.pone.0003046
- Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., Schroeder, C.E., & Foxe, J. J. (2005). Grabbing your ear: Rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex*, 15(7), 963–974.
- Ngo, M. K., & Spence, C. (2010). Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Attention, Perception, & Psychophysics*, 72(6), 1654–1665. doi:10.3758/APP.72.6.1654
- Nodine, C. F., & Kundel, H. L. (1987). Using eye movements to study visual search and to improve tumor detection. *Radiographics: a Review Publication of the Radiological Society of North America, Inc*, 7(6), 1241–1250.
- Olivers, C. N. L., & van der Burg, E. (2008). Bleeping you out of the blink: Sound saves vision from oblivion. *Brain Research*, 1242, 191–199. doi:10.1016/j.brainres.2008.01.070
- Orchard-Mills, E., Van der Burg, E., & Alais, D. (2013). Amplitude-modulated auditory stimuli influence selection of visual spatial frequencies. *Journal of Vision*, 13(3), 6. doi:10.1167/13.3.6. 1–17.
- Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual review of neuroscience*, 28, 377–401. doi:10.1146/annurev.neuro.27.070203.144216
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbi, M. I., Wu, W.-H. C., Cohen, L., ... Haxby, J. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proc Natl Acad Sci U S A*, 101(15), 5658–5663. doi:10.1073/pnas.0400707101
- Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. L. (2008). Haptic pop-out in a hand sweep. *Acta Psychologica*, 128(2), 368–377. doi:10.1016/j.actpsy.2008.03.011
- Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1), 3–25.
- Rudmann, D. S., & Strybel, T. Z. (1999). Auditory spatial facilitation of visual search performance: Effect of cue precision and distractor density. *Human Factors*, 41(1), 146–160. doi:10.1518/001872099779577354
- Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M. P., Dold, G., & Hallett, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, 380(6574), 526–528. doi:10.1038/380526a0
- Santangelo, V., Van der Lubbe, R. H. J., Belardinelli, M. O., & Postma, A. (2006). Spatial attention triggered by unimodal, crossmodal, and bimodal exogenous cues: A comparison of reflexive orienting mechanisms. *Experimental Brain Research*, 173(1), 40–48.
- Sathian, K., & Zangaladze, A. (2002). Feeling with the mind's eye: Contribution of visual cortex to tactile perception. *Behavioural Brain Research*, 135(1–2), 127–132.
- Sathian, K., Zangaladze, A., Epstein, C. M., & Grafton, S. T. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature*, 401(6753), 587–590. doi:10.1038/44139
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, 408(6814), 788. doi:10.1038/35048669
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA, USA: MIT Press.
- Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb Cortex*, 17(3), 679–690. doi:10.1093/cercor/bhk016
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9), 400–410. doi:10.1016/j.tics.2010.06.008
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, 135(2), 77–99. doi:10.1016/j.actpsy.2010.02.006
- Theeuwes, J., & Van der Burg, E. (2007). The role of spatial and nonspatial information in visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 1335–1351. doi:10.1037/0096-1523.33.6.1335
- Theeuwes, J., & van der Burg, E. (2008). The role of cueing in attentional capture. *Visual Cognition*, 16(2–3), 232–247. doi:10.1080/13506280701462525
- Theeuwes, J., & Van der Burg, E. (2011). On the limits of top-down control of visual selection. *Attention, Perception, & Psychophysics*, 73(7), 2092–2103. doi:10.3758/s13414-011-0176-9
- Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition*, 14(4–8), 466–489. doi:10.1080/13506280500195110
- Van der Burg, E., Cass, J., Olivers, C. N., Theeuwes, J., & Alais, D. (2010). Efficient visual search from synchronized auditory signals

- requires transient audiovisual events. *PLoS One*, 5(5), e10664. doi:10.1371/journal.pone.0010664
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008a). Audiovisual events capture attention: Evidence from temporal order judgments. *Journal of Vision*, 8(5), 2. doi:10.1167/8.5.2. 1–10.
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008b). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5), 1053–1065. doi:10.1037/0096-1523.34.5.1053
- Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2009). Poke and pop: Tactile-visual synchrony increases visual saliency. *Neuroscience Letters*, 450(1), 60–64. doi:10.1016/j.neulet.2008.11.002
- van Polanen, V., Tiest, W. M. B., & Kappers, A. M. L. (2012). Haptic pop-out of movable stimuli. *Attention, Perception, & Psychophysics*, 74(1), 204–215. doi:10.3758/s13414-011-0216-5
- Violenteyev, A., Shimojo, S., & Shams, L. (2005). Touch-induced visual illusion. *Neuroreport*, 16(10), 1107.
- Wang, M.-J. J., Lin, S.-C., & Drury, C. G. (1997). Training for strategy in visual search. *International Journal of Industrial Ergonomics*, 20(2), 101–108. doi:10.1016/S0169-8141(96)00043-1
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6), 495–501.
- Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'Connell, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1), 34–49. doi:10.1037/0096-1523.18.1.34
- Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, 44(12), 1411–1426. doi:10.1016/j.visres.2003.11.024
- Zannoli, M., Cass, J., Mamassian, P., & Alais, D. (2012). Synchronized audio-visual transients drive efficient visual search for motion-in-depth. *PLoS One*, 7(5), e37190. doi:10.1371/journal.pone.0037190