VISION IS SUPERIOR TO TOUCH IN SHAPE PERCEPTION EVEN WITH EQUIVALENT PERIPHERAL INPUT

Yoonju Cho,1,2 J.C. Craig,4 S.S. Hsiao,1,3† S. J. Bensmaia5¥

1 Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, MD 21218
2 Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218
3 Department of Neuroscience, Johns Hopkins University, Baltimore, MD 21218
4 Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405
5 Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637
†Deceased.
¥Email: sliman@uchicago.edu; Address: 1027 E. 57th St, Chicago, IL, 60637

ACKNOWLEDGMENTS

Steven Hsiao played an important role in the design of the experiments described in this paper and is therefore included as a co-author despite his passing in June of 2014. This work was supported by NIH grant NS 018787 and NSF grant IOS 1150209.
ABSTRACT

Results from previous studies suggest that two-dimensional spatial patterns are processed similarly in vision and touch when the patterns are equated for effective size or when visual ‘stimuli are blurred to mimic the spatial filtering of the skin. In the present study, we measured subjects’ ability to perceive the shape of familiar and unfamiliar visual and tactile patterns to compare form processing in the two modalities. As had been previously done, the two-dimensional tactile and visual patterns were adjusted in size to stimulate an equivalent number of receptors in the two modalities. We also distorted the visual patterns using a filter that accurately mimics the spatial filtering effected by the skin to further equate the peripheral images in the two modalities. We found that vision consistently outperformed touch regardless of the precise perceptual task and of how familiar the patterns were. Based on an examination of both the earlier and current data, we conclude that visual processing of both familiar and unfamiliar two-dimensional patterns is superior to its tactile counterpart except under very restricted conditions.
INTRODUCTION

Proficient Braille users are capable of reading a hundred words per minute through palpation of spatially patterned embossed dots. Similarly, letters generated by an optical-to-tactile converter (Optacon), a reading device for the blind, can be read by a skilled user at rates of 40 to 60 words per minute or more. Even naïve subjects can identify letters of the alphabet tactually with reasonable accuracy in laboratory settings (Craig 1979; 1981a; b; 1976; Loomis 1980; 1981a; 1974; 1985; Loomis and Apkarian-Steilau 1976; Phillips et al. 1983; Vega-Bermudez et al. 1991). Results from previous studies indicate that tactile spatial processing is quite good and may even be as good as its visual counterpart when adjustments are made for the obvious differences at the sensory periphery. In these studies, spatial patterns were presented visually and tactually and the visual pattern was manipulated to compensate for these differences. In one series of studies, visual letters were blurred to mimic the low-pass filtering of the skin, which eliminates the higher spatial frequencies (Loomis 1982; Loomis, 1990). With sufficient blurring, visual performance levels matched their tactile counterparts, and the confusions between letters were similar in vision and touch.

In another study, the sizes of tactile and visual letters were adjusted so as to equate the number of receptors activated in the two modalities (Phillips et al. 1983). In the finger pad, the afferents thought to convey spatial information at the highest resolution, slowly adapting type 1 (SA1) fibers (Johnson and Phillips 1981; Phillips and Johnson 1981), innervate the skin with a density of about 100 units per square centimeter, spaced about 1 mm apart (Johansson and Vallbo 1979). At the fovea, the equivalent area spans about five minutes of arc (Curcio et al. 1990). When both the tactile and visual letters spanned about 5 receptors (5 mm, 2.5 minutes of arc, respectively), letter identification performance was similar for the two modalities, about 50% correct. Furthermore, the patterns of confusions were similar in the two modalities, as had been found in Loomis’s studies. Together, these studies supported the hypothesis that visual and tactile form processing rely on analogous neural mechanisms.

Although the consensus favors the idea that visual and tactile processing of spatial patterns is similar when differences at the periphery are accounted for, there are data that point to differences in central processing between the senses. For example, Loomis (1993) found differences in the effect of surrounds on the perception of visual and tactile spatial patterns. More directly relevant to the present study, Loomis noted that with larger-sized patterns, vision consistently outperformed touch (Loomis, 1990).

One aim of the present study was to extend the comparisons of visual and tactile pattern perception to unfamiliar patterns. In previous studies of tactile pattern identification, almost without exception, subjects are presented with patterns such as letters, with which they are already familiar visually, or with patterns such as Braille or Braille-like patterns, on which sighted subjects receive visual training. Subjects may thus have learned to associate a tactile pattern with its corresponding visual representation. One aim of the present study was to examine tactile pattern perception when subjects lack readily available visual representations of the patterns.

The second aim of the study was to equate the peripheral images in the tactile and visual modality using a principled approach. First, we equated the number of activated receptors (as had Phillips and colleagues). Second, we filtered the visual image using a model of skin
mechanics that precisely reproduces the spatial filtering properties of the skin (Sripati et al. 2006). The present study consisted of a series of three experiments. In these experiments, subjects were presented with spatial patterns, either tactually or visually. In Experiments 1 and 3, the task was to identify the pattern. In Experiment 2 the task was to discriminate between two patterns.

**METHODS**

**Tactile stimulus presentation**

Because one of the aims of this study was to compare tactile and visual processing, we sought to equate the stimulus presentation in the two modalities to the extent possible. The tactile stimuli were presented through a tactile stimulator, consisting of 400 probes – each under independent computer control – arrayed in a 20 x 20 matrix, spanning a 1 cm x 1 cm area (Killebrew et al. 2007). This array is capable of delivering complex spatiotemporal patterns to the skin, as a computer monitor does to the retina. The subject’s left index finger was pressed against the probes using a restraint that allowed for accurate, repeatable finger positioning on the stimulator surface and a constant contact force. The density of the probes is greater than the density at which SA1 afferents – primarily responsible for conveying fine spatial information – innervate the fingertip skin. The probe density leads to a smooth percept despite the intrinsic discreteness of the array (Bensmaia et al. 2006a; Bensmaia et al. 2006b). The patterns were generated by indenting the probes 500 µm into the skin. Unless otherwise noted, patterns were presented for 500 ms after which the probes were withdrawn from the skin. The size of the largest tactile stimuli was 100 mm² (10 mm x10 mm) and thus spanned the receptive fields of approximately one hundred SA1 afferents (cf. Johansson and Vallbo 1979).

As mentioned above, we sought to equate conditions between vision and touch to the extent possible. Both the finger holder and the brief presentation time, 500 ms, restricted the amount of active exploration by touch of the pattern. However, we could not completely eliminate small eye movements, so is it possible that allowing visual but not tactile relative motion might provide an advantage for vision? With the small size (5 minutes of visual angle) and brief presentation times (500 ms), active visual exploration of the patterns is unlikely. Although there are microsaccades in the range of 5 minutes of visual angle, their main function seems to be to prevent perceptual fading (Martinez-Conde, Otero-Millan, and Macknik, 2013), which is not a significant problem with brief presentation times. There is the somewhat more complicated issue of the effectiveness of relative motion between a tactile pattern and the finger pad. While relative motion is critical for perceiving texture (see e.g. Hollins & Risner, 2000), the advantage is less clear for perceiving shape. Craig (1981b) found that patterns were more easily identified when presented briefly without relative motion than when scanned across the finger pad. One might hypothesize that subjects in the visual tasks might benefit from active exploration of the stimulus, denied to the subjects in the tactile tasks. However, Vega-Bermudez, Johnson, and Hsiao (1991) found no advantage for active exploration of tactile spatial patterns versus passive reception of the pattern. Loomis (1981a) did, however, find better performance with active exploration. Thus, it is not clear that relative motion or active exploration would improve tactile pattern recognition. The lack of clear-cut consensus on the advantage of motion and/or active exploration plus the requirement to precisely control the pattern presentation times, to have highly repeatable stimuli, and to have comparable visual and tactile conditions led to our decision to restrict finger movement.
Visual stimulus presentation

Visual stimuli were presented with an IBM ThinkPad W500 laptop. To achieve equivalence with the tactile modality, we adjusted the size of the visual patterns to stimulate the same number of receptors as the tactile patterns. The height of the visual stimulus was adjusted to span about 5 minutes of arc or the receptive fields of about one hundred cones (in the fovea, cf. Curcio et al. 1990). To achieve such small visual stimuli, we used patterns that were 3.4 mm x 3.4 mm on the laptop monitor with a 1440 x 900 resolution and with the monitor positioned 2.4 m from the subjects. Patterns were at high contrast (black on white) and the monitor was set to maximum brightness.

To further equate the visual and tactile patterns, we passed the visual patterns through a filter for some of the comparisons with tactile performance. Specifically, this filter mimics the propagation through the tissue of forces applied to the surface of the skin and can be used to predict both the responses of mechanoreceptive afferents to (Sripati et al. 2006) and the perception of (Craig et al. 2008; Pei et al. 2008) tactile spatial patterns. In brief, the model first computes the forces that the stimulus exerts at the surface of the skin. Then, it invokes continuum mechanics to estimate how forces at the skin’s surface would propagate through the tissue (Timoshenko and Goodier, 1970) to the location of the receptor (at a depth of 700 µm for SA1 afferents, cf. Sripati et al., 2006), thereby causing a deformation of the receptor, which in turn depolarizes it. The skin tends to act as a low-pass filter but also increases the salience of certain features, such as corners, which further distorts the spatial configuration of the stimulus presented at the surface of the skin (Sripati et al., 2006). To apply this model, then, we computed the strain at the depth of Merkel receptors for each stimulus, then converted the pattern of strains into a gray scale with the maximum strain corresponding to black and zero strain corresponding to white.

Experiment 1: match-to-sample

Subjects were presented with a complex target pattern, either visually or tactually, and selected which of five visually presented patterns matched the target. There were 112 target patterns out of a pool of 731 possible patterns (see below). None of the subjects in this or subsequent experiments had prior experience reading Braille. All testing procedures were performed in compliance with the policies and procedures of the Institutional Review Board for Human Use of the University of Chicago.

Subjects

Of the 56 subjects who participated in this experiment, 20 (7 females) participated in the experiment with tactile targets, 19 (3 females) in the experiment with unfiltered visual targets, and 17 (7 females) in the experiment with the filtered visual targets. Subjects only participated in one experiment to eliminate the possibility that they might learn the spatial patterns. All subjects had normal or corrected to normal vision and reported no tactile abnormalities. Subjects ranged in age from 18 to 32 years old.

Stimuli

The 112 target patterns were generated from seven sets of stimuli (Figure 1A). The first four sets were organized in a quadrant configuration, in which the spatial complexity of the stimuli varied
systematically. Three of the sets were bar patterns, consisting of two, three, or four bars, each
bar at one of four orientations (varying from 0 to 135° in 45° increments) and placed in one of
four quadrants (0 degrees indicates an orientation perpendicular to the long axis of the finger).
The fourth set (DB) consisted of combinations of dots and bars placed in three of four
quadrants; the bars were at one of four orientations (0 to 135° in 45° increments) and each dot
was centered in its quadrant. The fifth set (CS) consisted of combinations of curved lines (half-
spheres) and straight lines. The sixth set comprised Braille letters, the seventh set, letters of the
English alphabet. Braille characters were modified to fit the stimulus area and Helvetica font was
used for the letter stimuli. Importantly, some stimuli were organized in a quadrant structure
(bars and DB, e.g.) whereas others were not. Therefore, subjects could not encode all of the
stimuli using a quadrant-based mnemonic, for example by sequentially attending to each
quadrant and verbally encoding the local stimulus component. Indeed, this strategy would not
work with a subset of components, which, we surmised, would encourage subjects to focus on
the whole pattern. Each target category included 16 patterns, yielding a total of 112 target
patterns. The matching stimuli were generated from the seven sets of patterns in the same way
as were the target patterns. A pool of 731 patterns was thus generated: 112 target patterns and
619 foils.

As mentioned above, tactile targets spanned an area of 1 cm x 1 cm and visual targets spanned
5 min of arc. Visual targets were presented in two different ways: images that were equated to
their tactile counterparts only in size (“unfiltered”) and the same-sized images passed through a
filter that mimics skin mechanics (“filtered”).

**Procedure**

Subjects were tested individually either with tactile target patterns or with visual target
patterns. Subjects in the tactile condition were seated with their left index finger pad in contact
with the tactile array and with a visual monitor located 2.4 m away. Note that, on visual trials,
there were two monitors: One located 2.4 m away to display the target, the other located, also
2.4 m away, to display the matches. Each visual trial began with a 500-ms fixation period
indicated by a fixation cross at the center of the visual display. The fixation period was followed
by the presentation of the target pattern, either visual or tactile, for 500ms, which in turn was
followed by five choice patterns. One of the five patterns matched the target pattern. The
choice patterns subtended a visual angle of 10 minutes of arc and were arranged in a row,
spaced approximately 15 minutes of arc apart. Given their size, the shapes of the choice
patterns were easily perceptible. The five patterns remained on the screen until the subject
responded using a computer mouse. No feedback was provided. Each response was followed by
a 1-s period before the next trial began.

There were two types of trials: On half of the trials, so called “within-type” trials, all of the foils
were from the same set of patterns as the target, forcing the subjects to choose among similar
patterns. For example, if the target pattern consisted of two line segments, then the five choice
patterns also consisted of two-line patterns. On the other half of the trials, “random-type” trials,
the four foils were randomly selected from one of the remaining six categories. With these
random-type trials, the choices were generally quite different from the target pattern. By having
two types of matching trials, within and random, two levels of difficulty were created (Figure
1B). Each target was presented 5 times in pseudorandom order. There were 112 trials per block
and 5 blocks of trials for a total of 560 trials per condition.
Experiment 2: same-different

As discussed in more depth below, the objective of this experiment was to ascertain whether any differences in performance between the visual and tactile conditions observed in Experiment 1 might be attributable to the fact that the matching patterns were always presented visually.

Subjects

Fifteen subjects participated in the experiment with the unfiltered visual stimuli and tactile stimuli. Eight subjects (four females) participated in the cross-modal experiment: visual-tactile or tactile-visual. Ten subjects were tested in the within modality discrimination (seven females). Three subjects participated in both paradigms, first in the cross-modal then the unimodal paradigm; their performance did not differ from that of the other subjects. Eight subjects (four females) participated in experiments with filtered visual targets. One of them was tested on both non-filtered and filtered images but she exhibited the same pattern of performance as the subjects who were tested on one or the other.

Stimuli

We delivered bar patterns that differed from one another in the location and orientation of the bars. Patterns consisted of one to four bars, each at one of four orientations (0 to 135 in 45 increments). In the one-bar patterns, the bar was located at the center of the stimulus area (Figure 1C). In the two-bar patterns, if both bars were in one row or one column, the pattern was centered on the display (horizontally or vertically, respectively) (see Figure 1B). One way we generated a choice pattern that differed from the target was by rotating one of the bars by 90°; the large rotation ensured that the difference in orientation was easily perceptible (Bensmaia et al. 2008a; Bensmaia et al. 2008b). For the two- and three-bar patterns, another set of “different” patterns was generated by relocating one of the bar components to a new quadrant, without changing the orientation of the altered bar. As a result, individual one- and four-bar patterns had only one type of “different” pattern paired with it, namely involving a change in orientation. Each two- and three-bar patterns could be paired with two types of possible “different” patterns, one involving a change in orientation, the other involving a change in the position of one of the bars. There was a total of 36 patterns. As in Experiment 1, tactile targets spanned an area of 1 cm x 1 cm and visual targets spanned 5 min of arc; visual targets were presented both unfiltered and filtered in separate conditions.

Procedure

Separate blocks of trials (240 each) were run for each condition, tactile-tactile (t-t), visual-visual (v-v), tactile-visual (t-v), visual-tactile (v-t). On each trial, subjects were presented two patterns, each lasting 500 ms, separated by a 1-s interstimulus interval. As noted, with the two- and three-bar patterns, “different” patterns differed from the “same” pattern in two ways, either by the orientation of a bar or by the quadrant in which the bar appeared. The proportion of “different” trials for two- and three-bar patterns was approximately 2/3 to accommodate the two different types of transformations (orientation and location). Combining the percentage of “different” trials from the one- and four-bar trials with the two- and three-bar trials yielded an overall percentage of “different” trials of 60%. Subjects were not made aware of this and
received no trial-by-trial feedback. Subjects responded, “same” or “different” by pressing one of two keys.

**Experiment 3: letter identification**

The objective of this experiment was to assess subjects’ performance on tasks that have been extensively used in the past to compare shape processing in vision and touch, namely letter identification. We also wished to assess the effect of stimulus duration on identification performance to assess whether the short durations used in Experiments 1 and 2 may have affected the subjects’ performance levels.

**Subjects**

A total of 21 subjects participated in this portion of the study. Nine (six females) subjects were involved in the tactile part while twelve participated in the visual experiment, six (five females) on regular visual letters and six (five females) on filtered visual letters. One subject completed both visual and tactile tasks but exhibited comparable performance as subjects who were tested in only one modality.

**Stimuli**

Target stimuli consisted of the 26 letters of the English alphabet, presented in Helvetica font. Tactile patterns were 1 cm high, and visual ones spanned 5 min of arc in vertical extent.

**Procedure**

Subjects were presented with one of 26 letters of the alphabet, either tactually or visually, and their task was to identify the letter. Each letter was presented once in random order, the subject identified the letter, and no feedback about performance was provided. The duration of letter presentation (visual and tactile) was 0.5, 1, 2, or 4 seconds, and varied randomly from trial to trial. At the beginning of each visual trial, there was a 500-ms fixation point that was followed by the presentation of the target letter. Subjects indicated the target letter by pressing the corresponding key on a keyboard. After indicating the target letter, subjects pressed the enter key which recorded their response and initiated the next trial after a 500-ms pause. There were 104 trials per block and 5 blocks. The first run was for practice to allow subjects to familiarize themselves with the task.

**RESULTS**

**Experiment 1: match-to-sample**

First, subjects performed better on random-type trials than on within-type trials as might be expected given that the five matching patterns on within-type trials tended to be more similar to the target than did those on random-type trials (Geldard and Sherrick 1965; Horner 1991). Second, although the two modalities were matched in terms of number of receptors activated and spatial filtering, visual performance was superior to tactile performance with every type of target (Kruskal-Wallis test: $H=123.5, p<0.0001$)(Figure 2). Third, as expected, filtering the visual patterns reduced performance (Kruskal-Wallis test: $H=50.5, p<0.001$) but did not reduce it enough to match tactile performance, particularly for within-type trials (Kruskal-Wallis test: $H=50.5, p<0.001$) but did not reduce it enough to match tactile performance, particularly for within-type trials (Kruskal-Wallis test:
However, although performance in the filtered visual condition was superior to that in the tactile condition, the pattern of tactile performance across target types was more similar to the filtered visual than it was to the unfiltered visual performance (Figure 3). In summary, filtering the visual image to more closely match its tactile counterpart rendered performance more similar in the two modalities but did not completely eliminate the difference between tactile and visual form processing.

The only patterns that subjects were familiar with were the letters. The relatively poor performance with these patterns suggests that previous visual experience with these patterns provided little advantage in matching the tactile patterns. Context might be playing a role here in that the letters were part of a larger pattern set: there were six times as many non-letter patterns as letter patterns. Subjects might not be able to use top-down processing for letters – e.g., a template matching strategy – when these are interleaved with much more frequent non-letter patterns.

Experiment 2: same-different

An inherent imbalance exists between the tactile and visual conditions in Experiment 1 and in virtually all of the previous tactile pattern identification studies, namely that subjects select their responses from a set of visual representations, whether the target is visual or tactile. Thus, the visual conditions involve the matching of one visual pattern to another whereas, in the tactile conditions, a tactile pattern must be matched to a visual one. In Experiment 2, then, we sought to eliminate this imbalance in the comparison of visual and tactile pattern processing. To this end, we had subjects perform a ‘same-different’ task that has been fruitfully applied to the study of 3-D form perception (Craddock and Lawson 2009). On each trial, subjects were presented with a visual or tactile pattern followed by either a visual or tactile pattern. On half the trials, the second pattern was the same as the first and on half the trials they were different. Patterns were presented in four conditions, each involving a different combination of the tactile and visual modalities: tactile-tactile (t-t), visual-visual (v-v), tactile-visual (t-v), or visual-tactile (v-t). The subjects’ task was to determine whether the two patterns were the same or different, regardless of modality.

In this experiment, we wished to address two questions. First, would the t-t condition (within modality comparison) result in performance levels equal to that of the v-v condition? Second, how would performance in the t-v condition (which matches the conditions in Experiment 1) compare with that in the t-t condition? That is, are subjects better at comparing two tactile patterns than they are at comparing a tactile and a visual one? Again, testing was conducted with both unfiltered and filtered visual patterns.

As shown in Figure 4, performance was found to be considerably better when subjects compared a tactile pattern with a second tactile pattern than when they compared a tactile pattern with a visual pattern; t-t performance was significantly better than t-v or v-t performance (Kruskal-Wallis test: $H=17.6, p<0.0001$ for t-v; $H=14.6, p<0.001$ for v-t). Although improved relative to the cross-modal comparisons, t-t performance still did not rise to the level of filtered or unfiltered v-v (Friedman test: $H=18.6, p<0.0001$ for unfiltered v-v; $H=10.8, p<0.001$ for filtered v-v).
If, as suggested by the results from Experiment 1, tactile pattern perception declines more rapidly with increasing pattern complexity than does visual performance, then one might expect to see t-t performance fall off more rapidly with increasing number of bars than v-v performance. Such was the case when comparing performance in the t-t and unfiltered v-v conditions. There was a significant effect of number of bars (repeated-measures ANOVA: $F(3,27)=37.1$, $p<0.0001$) and a significant interaction between modality (unfiltered v-v vs. t-t) and number of bars ($F(3,27)= 4.0$, $p<0.05$). When t-t performance was compared to filtered v-v performance, the significant effect of number of bars was also significant (repeated-measures ANOVA: $F(3,21)=66.9$, $p<0.0001$), as was that of modality (filtered v-v vs. t-t) ($F(1,7)=6.2$, $p<0.05$), but the interaction between modality (filtered v-v and t-t) and the number of bars ($F(3,21)=0.6$, $p>0.5$) was non-significant. This lack of interaction suggests that filtering the visual stimuli renders performance more vulnerable to increasing complexity, as is the case in the tactile condition. While filtering the visual patterns did reduce performance, the effect did not quite reach statistical significance (unfiltered v-v vs. filtered v-v, Kruskal-Wallis test: $H=2.4$, $p>0.1$). This result may seem inconsistent with the dramatic effects of filtering on visual performance in Experiment 1. However, in Experiment 1, filtered visual targets were compared to unfiltered ones, whereas, in Experiment 2, filtered visual targets were compared to filtered ones. Since subjects did not need to “reverse” the filter to respond in Experiment 2, there was less reduction in performance with filtered images relative to unfiltered results.

Experiment 3: letter identification

In Experiments 1 and 2, vision consistently outperformed touch in pattern identification and matching. As noted, several previous studies used raised letters as tactile patterns and arrived at the conclusion that tactile and visual performance was similar or, at least, could be equated. We considered several possibilities as to why, in the present study, tactile performance was seemingly poorer than that observed in previous studies. One possibility was that, because the letters were embedded with unfamiliar patterns, subjects in Experiment 1 could not use a template matching strategy in the tactile conditions, in contrast to previous studies. Another possibility is that the quality of the tactile stimuli was somehow poorer than that of the visual stimuli, and that this difference in quality drove the observed behavioral differences. To test these hypotheses, we measured subjects’ ability to identify letters with the 400-probe array under conditions that matched those in Experiment 1 and also conditions that matched those in earlier studies with embossed letters.

In Experiment 3, visual letters were presented as they had been in Experiment 1, both “unfiltered” and passed through the tactile filter (but always equated for the number of receptors activated). As noted, this filter mimics the non-linear transformation produced by the deformation of the skin by a mechanical stimulus and its effect on mechanoreceptors (Sripati et al. 2006). In previous studies comparing tactile and visual letter identification performance, subjects were often allowed more time to view the patterns (Loomis 1981a; b; Phillips et al. 1983). For this reason, we varied the duration of the presentation of the patterns as well.

As expected, performance was found to improve as duration increases, but did so only modestly (Figure 5). First, we gauged how well subjects performed relative to earlier measures with raised letters. In two studies in which letter height was comparable to that in the present study (10 mm), identification performance was also comparable: 55% correct for 9-mm letters (Loomis 1981a); 62% for 8-mm letters (Phillips et al. 1983) compared to 58% correct in the present
study. In short, presenting letters via the 400-probe array produces levels of performance similar to those obtained with embossed letters. Second, we found that, in the visual condition, filtering reduced performance (Kruskal-Wallis test: $H=24.9$, $p<0.0001$), as we had previously found. Furthermore, visual performance was considerably and unexpectedly better than tactile performance for both the unfiltered and filtered patterns (Kruskal-Wallis test: $H=42.6$, $p<0.0001$ for unfiltered visual; $H=40.3745$, $p<0.0001$ for filtered visual). This result was initially surprising given that the 400-probe stimulator produced letter identification performance as good as that obtained with embossed letters, and in light of the generally held view that, when spatial patterns are presented to the same number of receptors, performance for the two modalities is equivalent (Phillips et al. 1983).

However, a closer look at previously obtained functions relating performance to letter height reveals a likely explanation and a critical finding that alters the view of the two sensory systems. Letter identification performance improves as letters get larger for both vision and touch. However, the tactile and visual functions have different slopes. Performance with statically presented letters is only equivalent in vision and touch when the letter height is 4.5 mm and 2 min of visual arc, respectively (both theoretically exciting the same number of receptors). At greater letter heights, visual performance outpaces tactile performance. At 3 min of visual angle, letter identification is about 80% correct and does not appear to be leveling off, whereas with the (equivalent) 5.5-mm letters performance is only about 62% correct. The Phillips et al. data suggest that with letters of 10 min of visual angle, the size used in the present study, visual performance would likely reach better than 90% correct, similar to that found here (Figure 5).

The Loomis study (Loomis 1981a) shows a similar pattern of results: As letter height increases, the separation between visual and tactile performance increases. As the letter height increased from 7.2 to 8.6 mm, visual performance continued to increase, whereas tactile performance leveled off (or decreased slightly). In short, the superior performance with visual letters in Experiment 3 appears to be consistent with the results from Experiments 1 and 2.

**DISCUSSION**

**Superiority of vision over touch**

In Experiment 1, we found that visual performance was consistently superior to tactile performance in identifying unfamiliar patterns. In Experiment 2, we found that unimodal pattern discrimination was easier than its cross-modal counterpart. The fact that tactile-tactile performance was superior to tactile-visual performance suggests that the poorer tactile performance in Experiment 1 resulted in part from requiring subjects to match tactile patterns to visual ones. The visual-visual performance, however, was still superior to the tactile-tactile performance, despite the fact that cross-modal matching was no longer required. However, this difference was reduced when comparing visual-visual matching with filtered stimuli to its tactile-tactile counterpart, which suggests that the filtering of the skin contributes to the superiority of vision over touch. In Experiment 3, visual letter identification performance was found to be superior to its tactile counterpart. Furthermore, we found that the 400-probe array used in the present study produced levels of performance similar to those obtained in previous studies with embossed letters, so the effects observed in Experiments 1 and 2 cannot be attributed to poorer stimulus presentation in the tactile conditions.
Why, in nearly all cases, did the visual modality outperform the tactile modality? One possibility is that there was an error in equating the two modalities for number of receptors: the visual patterns might be larger than their tactile counterparts when measured in terms of number of stimulated receptors. Given that innervations densities were derived from published data, it is unlikely that the discrepancy was large, however. Furthermore, tactile results were similar to those obtained in a previous study (Phillips et al. 1983), as were the unfiltered visual results (Experiment 3). Based on the present results and this closer examination of earlier results, the current view that the tactile modality performs just as well as the visual modality may be incorrect; our conclusion is that under equivalent stimulus conditions visual pattern perception is generally superior to tactile pattern perception. Note that a similar conclusion was drawn when comparing visual to tactile orientation discrimination (Bensmaia et al. 2008b).

Effect of familiarity

In Experiments 1 and 2, we found that tactile performance with unfamiliar patterns was generally poorer than its visual counterpart, even when the size of the visual patterns was adjusted to equate the number of receptors for the two modalities and their spatial features filtered to mimic the spatial filtering at the somatosensory periphery. The initial conclusion that lack of familiarity might be the cause of poorer tactile performance was called into question by the results from Experiment 3: the visual modality outperformed the tactile modality even with familiar patterns. On a related note, the present results suggest that the availability of visual templates does not result in a major improvement in the tactile pattern perception. In the tactile condition of Experiment 3, subjects knew that they would be identifying letters but their performance in identifying them was equivalent to that in Experiment 1, in which letters were interleaved with other patterns.

Effect of filtering

Transforming the visual image by filtering it through a skin-mechanics model so that it matched its tactile counterpart resulted in a significant decrement in performance. As expected, this decrement was particularly salient for more complex patterns (4-bar patterns vs. 2 bar patterns, for example, in the within-type conditions, Figure 2). Indeed, more complex patterns are more susceptible to distortion, as they comprise more features that can interact with one another. Filtering rendered visual performance qualitatively more similar to tactile performance: (A) the correlation between visual and tactile performance was greater for filtered than for unfiltered visual patterns; (B) Filtered visual performance was equally susceptible to increasing stimulus complexity as was tactile performance, while unfiltered visual performance was significantly less so. This suggests that the filtering produced by the skin is an important determinant of the superiority of vision over touch in shape perception.

Conclusions

Neural representations of shape have been shown to be remarkably analogous along the visual and somatosensory neuraxes. Indeed, in both modalities, spatial patterns are represented in the spatial pattern of activation of populations of receptors (in the skin and in the retina) which are then decomposed into oriented contours at the earliest stages of cortical processing (Bensmaia et al. 2008a). At higher cortical levels, tactile shape processing continues to bear strong similarities with its visual counterpart: feature selectivity becomes more complex, for example
consisting of both curvature and orientation tuning (Yau et al. 2013) and more invariant, for example exhibiting positional invariance for orientation selectivity (Fitzgerald et al. 2006). The superiority of vision over touch is not surprising when one considers their ecological functions. The function of shape processing in vision is to identify objects of arbitrary spatial complexity whereas that of touch is to manipulate them. While some information about shape is important to manipulate an object, manipulation likely requires a different kind of information than does identification. Thus, while the neural circuits responsible for tactile shape processing can be used for shape recognition, shape recognition is not their primary function. As a result, we have much more experience recognizing visual shapes than we do tactile ones (Loomis, 1990). The limited testing that our subjects received tactually could not begin to equate for a lifetime of visual experience with complex shape perception. A possible explanation, then, for the observed differences between vision and touch in shape perception is that the neural mechanisms of shape processing, while similar in the two modalities, are honed much more extensively in vision than in touch.
**Figure 1.** A| Two examples of each of the seven stimulus types: 2, 3, 4B: Two, three, four bar patterns; DB: dot and bar patterns; CS: Curved and straight line patterns; Br: Braille patterns; Lt: Letters. B| Example of a within-type and a random-type trial. In within-type trials, the target and all five matches are drawn from the same category (two bars in this case). On the Random-type trial, four choice patterns were from randomly-selected categories. The highlighted samples denote the matching stimulus C| Example targets from Experiment 2. Four types of targets consisted of one through four bars, presented unfiltered or filtered.

**Figure 2.** Performance on the match to sample task for touch, (black bars), and vision (unfiltered: white bars; filtered: gray bars). Each bar represents the mean performance across 1600 tactile trials, 1520 visual trials, and 1360 filtered visual trials. Error bars indicate the standard error of mean.

**Figure 3.** Visual vs. tactile performance for each stimulus type in Experiment 1. For within-type trials, the overall correlation between visual and tactile performance is 0.89 for unfiltered visual patterns and $r = 0.97$ for filtered visual patterns. For random type trials, visual and tactile performance is almost completely uncorrelated with unfiltered visual patterns ($r = -0.28$), but highly correlated for filtered ones ($r = 0.79$). Error bars indicate the standard error of mean.

**Figure 4.** Mean performance on the matching task of Experiment 2 for different pairs of modalities. Unimodal matching (reddish hues) yields better performance than does cross-modal matching. Error bars indicate the standard error of mean.

**Figure 5.** Mean performance on the letter identification task. Visual letters, equisized to the tactile letters, yield the best performance. When the skin filter is applied to visual letters, performance is lower than that for unfiltered letters but substantially higher than that for touch. Error bars indicate the standard error of mean.


### Table A

<table>
<thead>
<tr>
<th></th>
<th>2B</th>
<th>3B</th>
<th>4B</th>
<th>DB</th>
<th>CS</th>
<th>Br</th>
<th>Lt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
<td><img src="image4" alt="Image" /></td>
<td><img src="image5" alt="Image" /></td>
<td><img src="image6" alt="Image" /></td>
<td><img src="image7" alt="Image" /></td>
</tr>
</tbody>
</table>

### Table B

<table>
<thead>
<tr>
<th>Trial Type</th>
<th>Target</th>
<th>Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within-type</td>
<td><img src="image8" alt="Image" /></td>
<td><img src="image9" alt="Image" /> <img src="image10" alt="Image" /> <img src="image11" alt="Image" /> <img src="image12" alt="Image" /> <img src="image13" alt="Image" /> <img src="image14" alt="Image" /></td>
</tr>
<tr>
<td>Random-type</td>
<td><img src="image15" alt="Image" /></td>
<td><img src="image16" alt="Image" /> <img src="image17" alt="Image" /> <img src="image18" alt="Image" /> <img src="image19" alt="Image" /> <img src="image20" alt="Image" /> <img src="image21" alt="Image" /></td>
</tr>
</tbody>
</table>

### Table C

<table>
<thead>
<tr>
<th></th>
<th>Unfiltered</th>
<th>Filtered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1B</td>
<td><img src="image22" alt="Image" /></td>
<td><img src="image23" alt="Image" /></td>
</tr>
<tr>
<td>2B</td>
<td><img src="image24" alt="Image" /></td>
<td><img src="image25" alt="Image" /></td>
</tr>
<tr>
<td>3B</td>
<td><img src="image26" alt="Image" /></td>
<td><img src="image27" alt="Image" /></td>
</tr>
<tr>
<td>4B</td>
<td><img src="image28" alt="Image" /></td>
<td><img src="image29" alt="Image" /></td>
</tr>
</tbody>
</table>