Psychological Science
http://pss.sagepub.com/

"Please Tap the Shape, Anywhere You Like": Shape Skeletons in Human Vision Revealed by an Exceedingly Simple Measure

Chaz Firestone and Brian J. Scholl

Psychological Science published online 9 January 2014
DOI: 10.1177/0956797613507584

Research Article

Psychological Science
201X, Vol XX(X) 1–10
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797613507584
pss.sagepub.com

"Please Tap the Shape, Anywhere You Like": Shape Skeletons in Human Vision Revealed by an Exceedingly Simple Measure

Chaz Firestone and Brian J. Scholl
Yale University

Abstract

A major challenge for visual recognition is to describe shapes flexibly enough to allow generalization over different views. Computer vision models have championed a potential solution in medial-axis *shape skeletons*—hierarchically arranged geometric structures that are robust to deformations like bending and stretching. In the experiments reported here, we exploited an old, unheralded, and exceptionally simple paradigm to reveal the presence and nature of shape skeletons in human vision. When participants independently viewed a shape on a touch-sensitive tablet computer and simply tapped the shape anywhere they wished, the aggregated touches formed the shape's medial-axis skeleton. This pattern held across several shape variations, demonstrating profound and predictable influences of even subtle border perturbations and amodally filled-in regions. This phenomenon reveals novel properties of shape representation and demonstrates (in an unusually direct way) how deep and otherwise-hidden visual processes can directly control simple behaviors, even while observers are completely unaware of their existence.

Keywords

visual awareness, shape perception, sampling, attention

Received 4/16/13; Revision accepted 8/13/13

Among the most central challenges in the study of any mental process is to determine the format of its underlying representations. A case study arises from the problem of visual recognition. Views, positions, lighting, and other contextual factors ensure that objects and shapes rarely appear in just the same way across viewings, and so a challenge facing any prospective recognizer is to describe shapes flexibly enough to allow for generalizations over instances.

For decades, an influential solution championed by computer vision research has been to describe shapes according to a type of underlying *skeletal* representation, a leading candidate for which is a geometric structure known as the medial axis (Blum, 1973). The medial axis is the set of points having two or more closest points on the shape's perimeter, and so is arranged hierarchically: Parent branches coding global features of the shape sprout increasingly minor offshoots coding more local features. This property may allow recognition systems to efficiently match shapes at whichever grain of resolution is appropriate for a given task. The medial axis is also robust to deformations like bending and stretching: Just as the bones in your hand remain connected to each other as you make a fist or wave "hello," medial branches preserve their configuration over similar transformations. The success of shape skeletons in computer vision models (e.g., Liu & Geiger, 1999; Siddiqi & Pizer, 2008; Siddiqi, Shokoufandeh, Dickinson, & Zucker, 1999; Zhu & Yuille, 1996) raises the intriguing possibility that human vision employs this shape-representation strategy as well (e.g., Kimia, 2003), scrutinizing forms as if to infer their "blueprints"—the interior structures that explain how shapes came to have the exterior features they do (Feldman & Singh, 2006).

The possibility of skeletal shape representations in human vision has nevertheless resisted direct empirical study. Despite suggestions of perceptual prioritization for

Corresponding Author:
Brian J. Scholl, Yale University–Psychology, Box 208205, New Haven, CT 06520-8205
E-mail: brian.scholl@yale.edu

shapes' centers (Melcher & Kowler, 1999; Vishwanath & Kowler, 2003) and major symmetry axes (Harrison & Feldman, 2009; Kovacs & Julesz, 1994), previous studies using contrast sensitivity measurements (Kovacs, Feher, & Julesz, 1998) and probe detection (Barenholtz & Feldman, 2003; Wang & Burbeck, 1998) have not been able (and have rarely tried) to isolate medial-axis skeletal representations as distinct from representations based only on global symmetry. And although shape skeletons have motivated creative investigations into higher-level processes like categorization (Wilder, Feldman, & Singh, 2011) and even aesthetic judgment (Palmer & Guidi, 2011; van Tonder, Lyons, & Ejima, 2002), the methods used have allowed only for highly indirect appreciation of how actual skeletal shape representations might look (see also Hung, Carlson, & Connor, 2012).

One unheralded study predating all of these suggested a much more direct method of exploring shape skeletons. When many individuals drew a single dot in a location of their choosing within a shape's boundary on a piece of paper, the aggregated dot placements resembled the shape's medial axis, as if each subject's chosen location were sampled from an internal skeletal representation (Psotka, 1978). No study since has adopted Psotka's paradigm, and this report is very rarely cited in contemporary literature on skeletal shape representation. (For example, of the 19 subsequent reports on shape representation by interior structure cited in the present article, only 2 cite Psotka's work—both referencing it in passing, and without mentioning the experimental method.)

We think this paradigm was ahead of its time, both methodologically and theoretically. Not only does it bear on whether shapes are represented skeletally, but also it could be used to explore critical questions about how sensitive the mind's skeleton-extracting computations are to various factors identified in the shape-representation literature. For example, introducing even a slight perturbation into a shape's contour has profound consequences for the shape's medial axis as geometrically defined, as the medial axis sprouts offshoot branches that may or may not perspicuously describe the shape. Much contemporary work has thus focused on when and how to "prune" these (potentially spurious) branches, a problem often seen as the central challenge for skeletal shape matching (e.g., Pizer, Siddiqi, Szekely, Damon, & Zucker, 2003; Shaked & Bruckstein, 1998). But the visual system must itself have achieved a workable solution, and a phenomenon like the one described by Psotka (1978) could shed light on human vision's pruning function. More foundationally, this phenomenon may offer an unusually direct window onto the nature of mental representations inaccessible by introspection.

The Current Study

The present research explored the utility of this striking phenomenon by focusing on new theoretical questions, including the influence of border perturbations, surface features, and amodally filled-in regions; the role of sampling processes in deploying skeletal shape representations; and the extent to which metacognitive processes have (or do not have) access to such representations even as they directly control behavior. We displayed single closed geometric shapes on a touch-sensitive tablet computer (Fig. 1) and instructed more than 1,000 participants (all pedestrians in New York City's Times Square) simply to touch the displayed shapes, anywhere they wished. Though uncommon, this task may be ideal for allowing implicit representations ubiquitous in everyday visual experience to control behavior, because that

Fig. 1. Illustration of the experimental procedure. Pedestrians in New York City's Times Square were tested in individual 5-s sessions, with instructions to simply touch the displayed shape, anywhere they wished.

behavior is so seemingly arbitrary and unthinking (a situation similar, perhaps, to how a ubiquitous process like spreading activation in memory is best assessed with unusual tasks such as stem completion). Each participant contributed only one data point (in a roughly 5-s "session") and did not see other participants' chosen locations. Of interest were the patterns of aggregated touches.

General Method for Experiments 1 Through 7

Participants

Each experiment was run outdoors in New York City's Times Square. In total, 1,480 pedestrians participated, each providing a single touch of a single shape in a single experiment. Separate groups of 200 people each participated in Experiments 1, 3, 4, and 5; 400 people participated in Experiment 2; and separate groups of 140 people each participated in Experiments 6 and 7.

Apparatus and stimuli

Stimuli were presented using custom software written in Python with the PsychoPy libraries (Peirce, 2007, 2008) and appeared on a tablet computer with a capacitive 25.7-cm-diagonal touch screen and 1,280 × 800 resolution. All stimuli were black outlines of 2–D geometric shapes (line thickness: 0.05 cm), filling between 11% and 32% of the screen's area depending on the shape. Shapes in Experiments 1 through 5 were presented in different random locations on the display across participants and appeared on a white background. Shapes in Experiments 6 and 7 were presented such that their top edges were located in the display's center and appeared on a 50%-gray background (but had a white interior). Unless otherwise noted, the shapes had the same aspect ratio as the tablet (8:5).

Procedure

The experimenter explained to participants that he was running "the world's fastest psychology experiment" and that participants should "pick a spot anywhere inside the shape, and just gently tap that spot." After each touch, a new shape (for the next participant) was randomly chosen from the stock of shapes in circulation. The touch screen was cleaned regularly to prevent the accumulation of fingerprints.

Analyses

Occasionally, participants misunderstood (or flouted) the instructions and touched the screen outside the boundary of the shape. Touches falling more than 30 pixels (0.5 cm) outside of a shape were excluded from further analysis (2.7% of touches in total). All figures, statistical tests, and reported sample sizes reflect postexclusion totals.

Experiments 1 and 2: Medial-Axis Skeletons Versus Global Symmetry

The shape in Experiment 1 was an isosceles triangle. As visual inspection reveals, aggregated touches for this shape (which appear "raw" in Fig. 2a and as a heat map in Fig. 2b) conformed to the triangle's medial axis, whereas touches to nonmedial regions (e.g., the bottom edge) were strikingly rare. On average, touches were only 0.39 cm (23.2 pixels) from the nearest medial-axis point; by contrast, when we generated 50,000 distributions of 200 randomly and uniformly sampled points from the triangle's area, the simulated touches in the best of all 50,000 distributions averaged a distance of 0.69 cm (40.6 pixels) from the medial axis. The skeleton as a whole was also well covered by touches: A randomly sampled medial-axis point was only 0.14 cm (8.3 pixels) from the nearest touch on average. This constituted early evidence that shape skeletons are computed in human vision and can directly guide simple behavioral responses.

However, this result fails to distinguish whether touches tend to follow medial axes or other axes of global symmetry (which overlap for triangles)—a problem that also limited previous attempts to study shape skeletons in human vision. Therefore, for Experiment 2, we used a rectangle, which has a medial axis that is distinguishable from its global diagonal-symmetry axes. Participants' touches again conformed to the medial axis (Fig. 2c). Touches averaged a distance of 0.50 cm (29.2 pixels) from the nearest medial point, whereas the best of 50,000 random distributions of touches contained simulated touches averaging a distance of 1.06 cm (62.5 pixels) from the medial axis. The medial axis was also well covered by touches: On average, a randomly sampled medial-axis point was only 0.16 cm (9.5 pixels) from the nearest touch. Moreover, touches conformed to the medial axis rather than to the diagonal-symmetry axes (Fig. 2d): Touches averaged a distance of 0.57 cm (33.3 pixels) from the nearest diagonal-axis point, which was significantly farther than the average distance from the medial axis, $t(399) = 2.72$, $p = .007$ (all tests two-tailed)—and this comparison underestimates the superior fit of the medial axis, because the diagonal-symmetry axes are 10% longer than the medial axis. Participants' touches thus appear to be drawn distinctly toward medial axes.

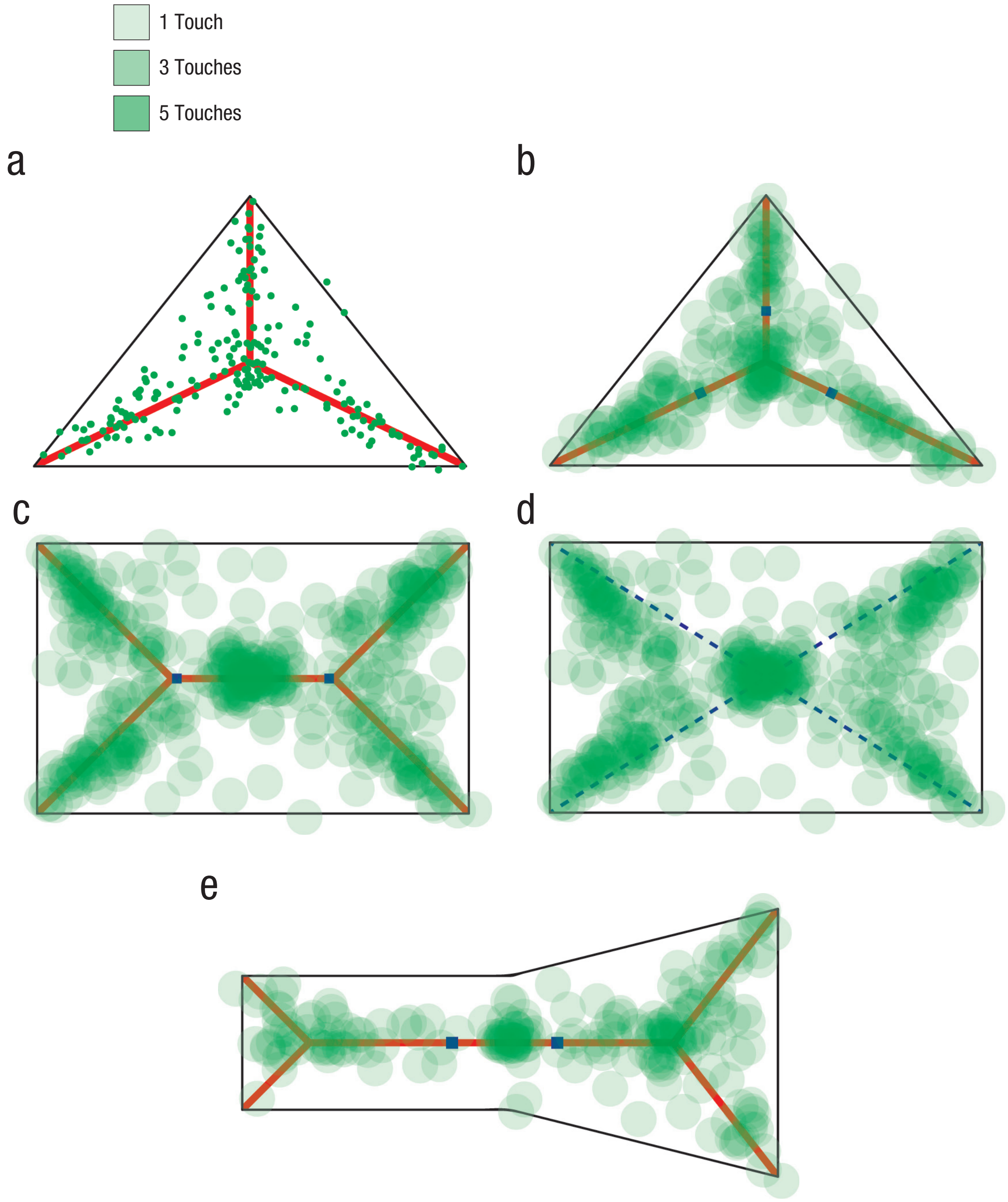

Fig. 2. Shapes, medial axes, and aggregated touches from Experiments 1 through 3. Solid red lines indicate medial-axis shape skeletons (not displayed to participants). Blue squares indicate the most sparsely touched regions on the medial axis as revealed by the searchlight procedure described in the text. For Experiment 1, aggregated touches are shown in raw form (a), with each dot depicting a touch location, and in a heat map (b). Two heat maps are shown for Experiment 2 to illustrate the distribution of touches relative to both (c) the medial axis and (d) the global diagonal-symmetry axes, indicated by the dashed blue lines (not displayed to participants). A heat map for aggregated touches in Experiment 3 is shown in (e). In the heat maps, each dot is 5 times larger than in (a) and is reduced to 15% opacity.

Experiment 3: Are All Medial-Axis Points Created Equal?

All points on the medial axis are "created equal," insofar as they are defined by the same geometric constraint (viz. having two or more closest points on the shape's perimeter). A close look at the data from Experiment 2, however, reveals two relatively touch-free regions flanking the rectangle's center along the medial axis (Fig. 2c). These regions can be identified quantitatively: A searchlight procedure considering a 50-pixel radius around each medial-axis point revealed that the two most sparsely covered points of relative equidistance from the center fall directly in the center of these subjective regions (and are depicted by the blue squares in Fig. 2c). Despite the relative centrality of these two points on the medial axis, only 2.8% of touches fell within those 50-pixel radii—fewer than what would be expected even if touches were randomly allocated over the figure (because the two circular regions cover 4.8% of the rectangle's total area).

What explains these sparsely covered regions? They could be caused by some intrinsic feature of skeletal shape representation. For example, they are extremely near the two axial junction points, and perhaps the visual system treats junctions differently than other areas. Alternatively, the sparse regions could be caused by a factor extrinsic to shape representation. For example, the two points are near the shape's center (2.15 cm, or 127 pixels, from it)—a point that may be prioritized for reasons entirely unrelated to the medial axis—and perhaps the center effectively "steals" touches for independent reasons. (Of course, these possibilities are perfectly confounded in the case of the triangle in Experiment 1.)

To distinguish these possibilities, in Experiment 3 we presented a guitarlike shape (depicted in Fig. 2e) consisting of a triangle fused with a rectangle. The guitar's wider aspect ratio makes its skeletal junctions farther from the shape's center, and its asymmetric contours make its skeletal junctions differentially distant from the center. This shape was presented half of the time as depicted in Figure 2e, and half of the time reflected about the *y*-axis (with the corresponding depicted touches also reflected).

The aggregated touches, also depicted in Figure 2e, again reliably fell on the medial axis. Touches averaged a distance of 0.41 cm (23.9 pixels) from the nearest medial-axis point, whereas simulated touches in the best of 50,000 random touch distributions averaged a distance of 0.74 cm (43.5 pixels) from the medial axis. The sparse regions on the shape flank the shape's center rather than tracking the skeletal junctions: The searchlight procedure described earlier identified a pair of points (enclosing only 4% of touches, and depicted in Fig. 2e by the blue squares) that were still near the center (and were in fact closer to it: 1.65 cm and 1.34 cm), but farther from the junctions (4.1 cm and 3.3 cm away).

Thus, the sparsely touched regions appear not to track anything specific to skeletal representations themselves. Instead, there is an independent bias for the center (Melcher & Kowler, 1999; Vishwanath & Kowler, 2003) that steals touches from surrounding regions.

Experiments 4 and 5: A Striking Influence of Relatively Minor Border Perturbations

The shapes tested in Experiments 1 through 3 had perfectly smooth contours, but real-world shapes often bear imperfections like bumps or dents (see Leyton, 1992). Much modeling work focuses on coping with such perturbations (e.g., Feldman & Singh, 2006; Siddiqi et al., 1999): If treated as signal, even a small notch profoundly rearranges the shape skeleton, which sprouts new branches that may not perspicuously describe the shape (see Fig. 3); if notches are dismissed as noise, then the skeleton remains undisturbed, but some shape information is lost. Of course, the visual system must itself have reached a solution to this problem—but this solution has never been explored experimentally. In Experiments 4 and 5, we therefore presented versions of the shapes from Experiments 1 and 2, introducing a concave, rectangular notch carved into one edge.

The aggregated touches conformed to these shapes' highly complex medial axes, extra branches and all

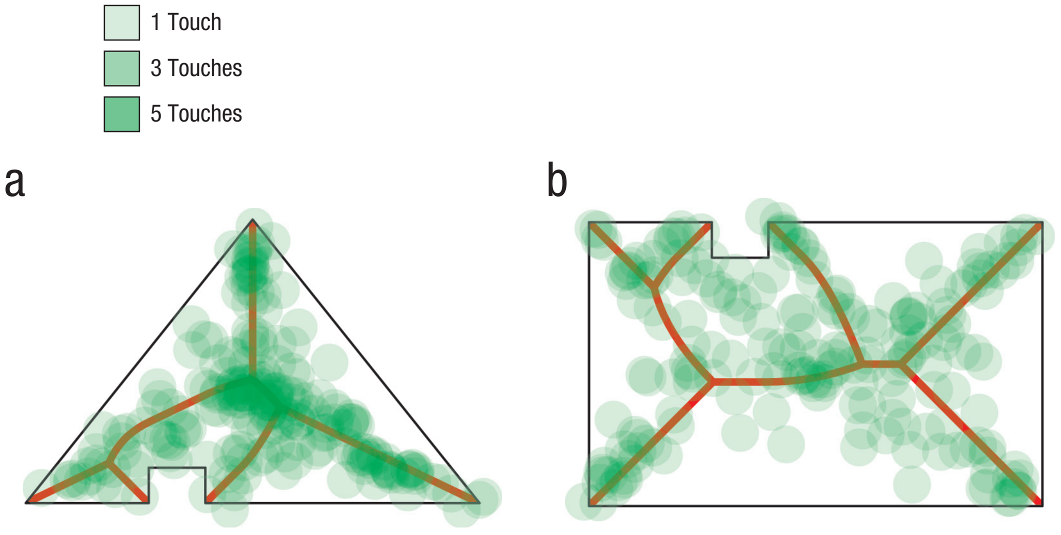

Fig. 3. Heat maps of aggregated touches for Experiments 4 (a) and 5 (b). Solid red lines indicate medial-axis shape skeletons.

(Fig. 3). Touches were better captured by the notched shapes' medial axes than by the unnotched shapes' medial axes, for both the triangle, $t(199) = 3.98$, $p < .001$, and the rectangle, $t(199) = 5.35$, $p < .001$. (Though the notched shapes' longer medial axes conferred a statistical advantage over the unnotched shapes' medial axes, neither notched shape's medial axis captured the unnotched shape's touches better than the unnotched shape's own medial axis did—triangle: $p = .22$; rectangle: $p = .23$). Additionally, a randomly sampled point (from 1,000 uniformly spaced samples) on a given notched shape's medial axis had a nearer nearest touch among the notched shape's touches than among the unnotched shape's touches—triangle: $t(999) = 9.74$, $p < .001$; rectangle: $t(999) = 9.87$, $p < .001$.[1] Evidently, even small border perturbations can figure in the computation of skeletal shape representations—a result that illustrates how this method can yield new discoveries about shape representation.

Experiments 6 and 7: Seeing Through Stickers—Shape Representation Per Se

In the examples presented so far—and in all previous discussions of the possibility of shape skeletons in human perception—the shapes' contours were fully visible. However, real-world shapes are often partially occluded. Using the tap-the-shape task, we were able to investigate for the first time the level at which shape skeletons are computed in visual processing: Are they constructed only from visible contours, or are they constructed from perceived shape based on higher-level factors?

We tested this using a shape that was amodally completed behind an occluding surface. In Experiment 6,

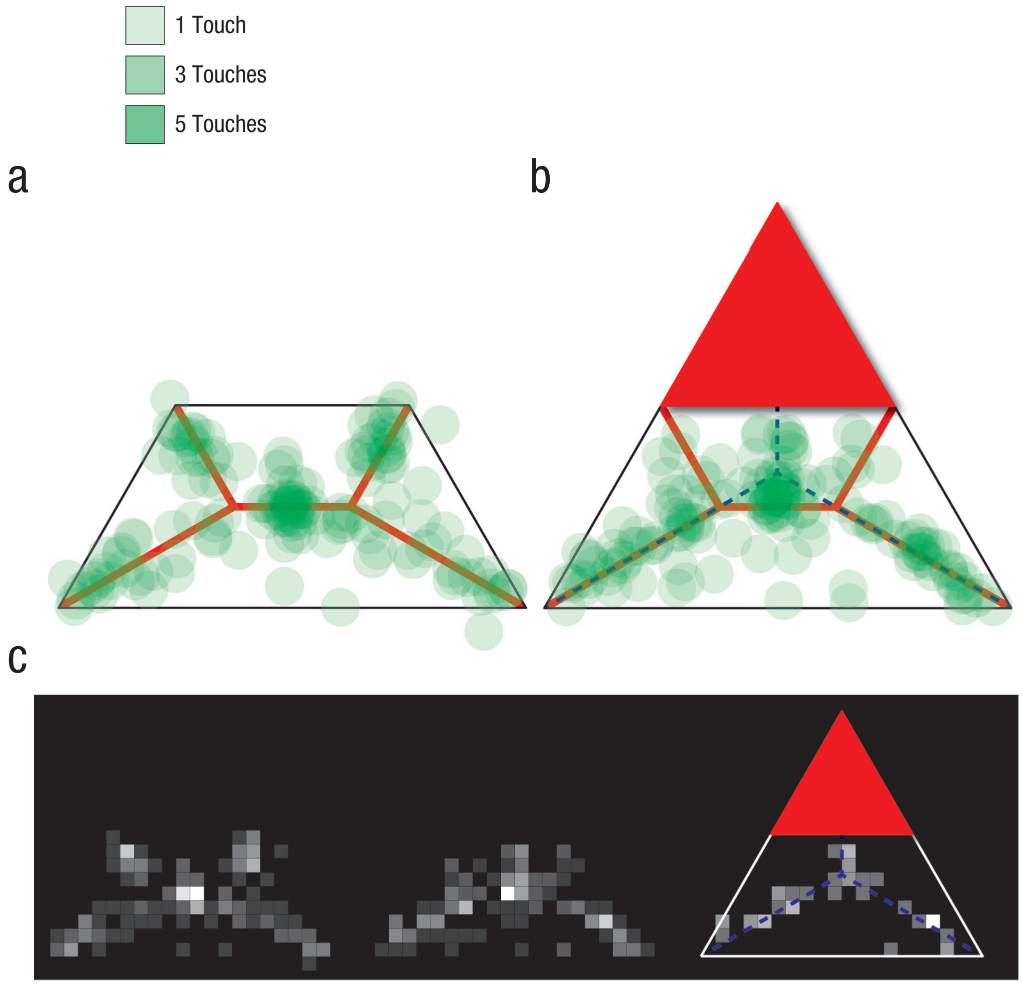

Fig. 4. Results from Experiments 6 and 7. The heat maps show aggregated touches from Experiments 6 (a) and 7 (b). Solid red lines indicate medial-axis skeletons for the trapezoids. In (b), the dashed blue line indicates the portion of the larger triangle's medial-axis skeleton contained in the trapezoidal area in which participants were instructed to touch, and the red triangle depicts the location and size of the sticker placed on the surface of the tablet computer's display. (The glossiness of the sticker made it unambiguously an occluding surface.) The diagrams in (c) show the matrices of local touch densities for Experiment 6 (left) and Experiment 7 (middle) and the results when the matrix for Experiment 6 was subtracted from that for Experiment 7 (right); brighter cells indicate more touches.

participants viewed a white trapezoid on a gray background (Fig. 4a) and were instructed to touch "anywhere inside the white area." In Experiment 7, we presented the very same shape, except that a bright red, triangular vinyl sticker (side length: 5 cm) was placed directly on the tablet's screen, just covering the trapezoid's uppermost edge, such that the trapezoid appeared to be the bottom half of a partially occluded triangle (Fig. 4b). (The sticker's gloss, material, and visible edges made it unambiguously an occluding surface, physically distinct from the display.)

The task in these two experiments was thus identical: to touch anywhere within a trapezoidal region. However, the trapezoid's appearance as part of a larger, filled-in triangle biased participants toward regions falling on the larger triangle's medial axis: Though touches of the stand-alone trapezoid (Fig. 4a) conformed to its medial axis (average touch-to-skeleton distance: 0.30 cm, or 17.9 pixels; coverage, or average skeleton-to-nearest-touch distance: 0.18 cm, or 10.9 pixels), touches of the trapezoid-plus-sticker (Fig. 4b) were concentrated on the skeletal branches shared by the filled-in triangle. The filled-in triangle's medial axis better captured touches of the trapezoid-plus-sticker than touches of the stand-alone trapezoid (average touch-to-skeleton distance: 0.41 cm, or 24.3 pixels, vs. 0.76 cm, or 45.1 pixels), $t(278) = 5.25$, $p < .001$. Moreover, touches of the stand-alone trapezoid better covered the upper branches of the trapezoid's skeleton (which do not appear on the triangle's skeleton) than did touches of the trapezoid-plus-sticker (coverage: 0.18 cm, or 10.7 pixels, vs. 0.30 cm, or 17.6 pixels), $t(999) = 16.35$, $p < .001$.

This pattern may be appreciated by subtracting a matrix of local touch densities for Experiment 6 from the same matrix for Experiment 7: The remaining cells (i.e., areas touched more in Experiment 7 than in Experiment 6) form all and only the branches making up the bottom portion of the filled-in triangle's medial axis (Fig. 4c). Participants were evidently drawn to the filled-in triangle's medial axis even when it was task irrelevant. Moreover, this result demonstrates that some figure-ground assignment precedes skeletal extraction and that shape skeletons can be influenced by higher-level shape properties.

This result also eliminates the possibility that a process unrelated to shape representation drove the previous shape-tapping results. For example, suppose participants began with no initial location preference (and would have touched a random point within the shape), but then avoided contours as they prepared and executed their response so as not to accidentally touch outside the designated area. If this contour-avoidance account were correct, then participants in Experiment 7 should also have (a) avoided the upper contour of the white trapezoidal area and (b) touched the upper skeletal "wings" as often as the rest of the trapezoid's skeleton—neither of which was the case (Fig. 4c). (The previous experiments also rule out this possibility, though. Contour avoidance straightforwardly entails that shapes' corners should be the last places participants would touch, but aggregated touches in all experiments reported here included many touches near corners.)

Experiment 8: Shape Skeletons Are Counterintuitive

Participants in the shape-tapping experiments appear to have sampled from internal skeletal shape representations in making their responses. Were they aware of this preference? Despite persistent claims that shape skeletons "intuitively" describe shapes (cf. "instant psychophysics"; Feldman & Singh, 2006, p. 18016), no study has investigated the intuitiveness of medial-axis shape descriptions as such (though some data may be indirectly relevant; Palmer & Guidi, 2011; van Tonder et al., 2002; Wilder et al., 2011). Experiment 8 is the first study of this question; we simply asked whether participants could predict the shape-tapping experiments' outcomes. Participants who had not completed any shape-tapping experiments read a description of the shape-tapping paradigm for the rectangle and selected which of nine illustrated touch distributions (e.g., random, clumped in the center, biased toward various global and local symmetry axes) they thought would obtain.

Method

Participants. One hundred naive participants were recruited online from Amazon Mechanical Turk. All participants gave informed consent and were given a small monetary reward for participating.

Stimuli and procedure. Participants read a description of the method of Experiment 2 and then were asked: "When the experimenter looks at where everyone touched the rectangle, what pattern of touches do you think he'll find? In other words, where do you think people will tend to touch the rectangle?" Participants then selected one among nine images of touch distributions (generated as depicted in Fig. 5 by randomly allocating dots along certain axes and adding a noise term), whose top-to-bottom order of appearance on a single Web page was randomized across participants.

Results and discussion

The responses (indicated numerically in Fig. 5) did not favor the medial axis, which was chosen at a below-chance rate (3% vs. 11.1%). Instead, participants favored the random distribution (33%) and the center-clumped

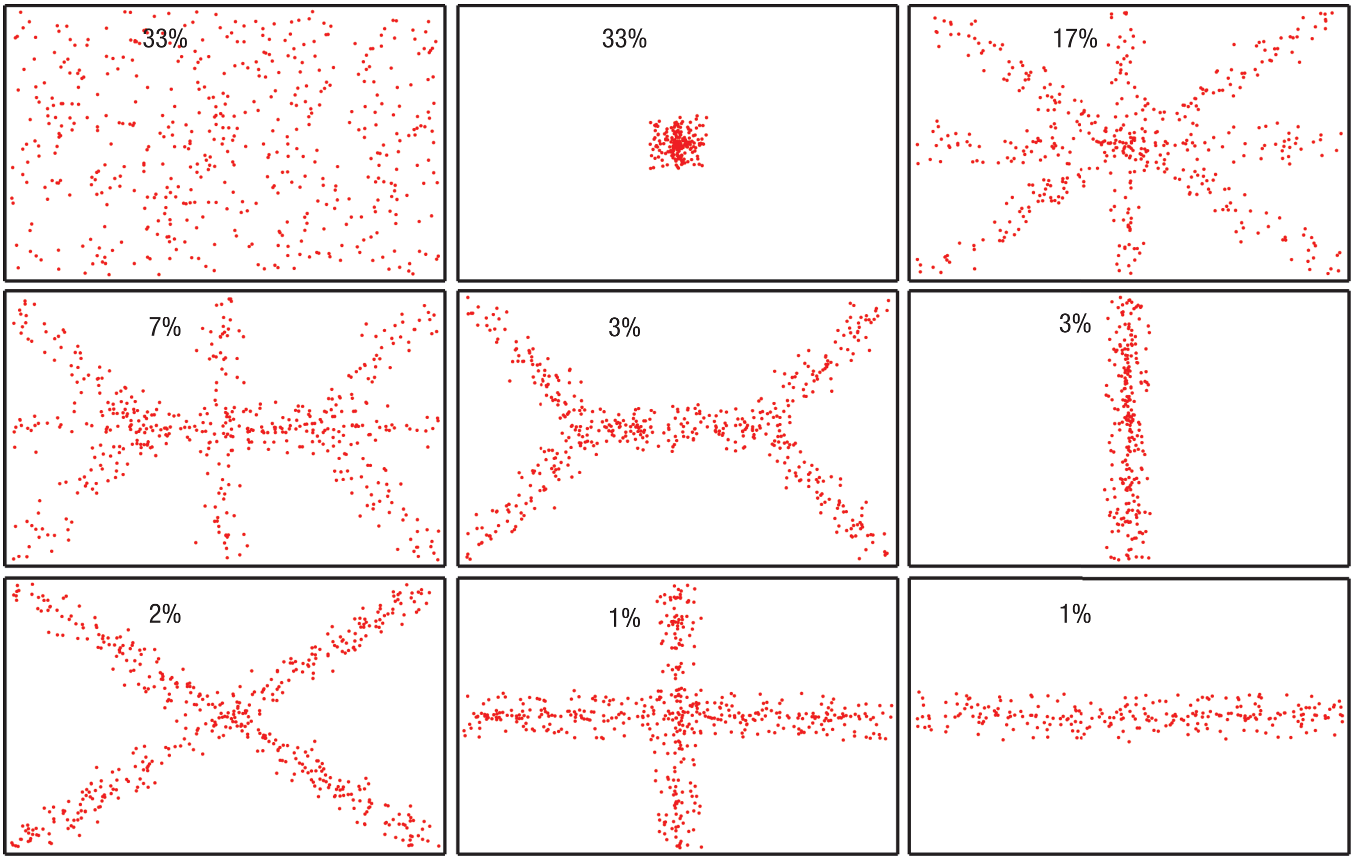

Fig. 5. The nine hypothetical distributions of touches used in Experiment 8. Naive participants selected among these distributions to indicate their prediction of the most likely outcome of Experiment 2. The number at the top of each panel (not displayed to participants) indicates the percentage of the participants who chose that option.

distribution (33%), followed by the global-symmetry distribution (17%). This suggests that medial axes are not especially intuitive shape descriptions, and perhaps even that participants were unaware of their own computation of skeletal shape representations (even though such representations guided their behavior in Experiments 1 through 7).

General Discussion

The seven shape-tapping experiments produced unusually direct evidence that human vision represents shapes in a skeletal format. When participants simply touched a shape anywhere they pleased, their chosen locations formed the shape's medial-axis skeleton. These findings complement a growing literature that has identified skeletal shape descriptions as useful codes of shape information for capacities like recognition (e.g., Burbeck & Pizer, 1995; Feldman & Singh, 2006; Kimia, 2003; Liu & Geiger, 1999; Siddiqi & Pizer, 2008; Siddiqi et al., 1999; Zhu & Yuille, 1996), adding particularly direct and easily appreciated evidence to the small body of empirical work on this question (e.g., Harrison & Feldman, 2009; Kovacs et al., 1998; Kovacs & Julesz, 1994; Wang & Burbeck, 1998; Wilder et al., 2011). The tap-the-shape task, being applicable to any shape, also allowed us to consider a variety of new questions, for example, concerning the influence of border perturbations (Experiments 4 and 5) and amodally filled-in regions (Experiments 6 and 7).

"Imaging" shape representations

However, we see a deeper worth for these results in the fact that obtaining them was even possible in the first place. The majority of mental representations posited in cognitive psychology (and especially in vision science) operate "under the hood," inaccessible even to the owners of the minds computing them. Indeed, such representations are frequently difficult to study for just this reason. Skeletal shape representations fall squarely into this category, lying far below the level of conscious report (see

Experiment 8). And this may help explain why what little evidence there is for skeletal representations in human vision has come from especially indirect measures (e.g., painstakingly constructed perceptual sensitivity maps revealed after thousands of trials per observer, as in Kovacs & Julesz, 1994). In contrast, the fast and simple tap-the-shape task manages to fish these representations out, almost as if by imaging them.

It is worth spelling out the nature of the inference from patterns of touches to conclusions about how shapes are represented. Why should the former indicate anything about the latter? We think the inference is simple: That the very same, intricate pattern is found in two ostensibly unrelated domains (i.e., many participants' shape touches and computer vision models of shape recognition) cries out for explanation, and a ready one is that there is some relation between the results. The medial axis is a very detailed, specific, and counterintuitive geometric structure. When that detailed structure is implicated in two different ways, there is an important sense in which it ought to be considered the *same thing*.

Sampling from skeletons

These findings also bear on other issues central to contemporary cognitive psychology. For example, even if the visual system computes skeletal shape representations, and even if such representations guide behavioral responses, touches in these experiments need not have widely covered the medial axis. In particular, participants might have computed a shape skeleton, but then touched a preferred location on the skeleton (e.g., the center of mass). In contrast, our results suggest that participants genuinely sampled over a distribution (of medial-axis points) in making their responses, a notion that has recently enjoyed popularity in cognitive psychology (e.g., Vul, Hanus, & Kanwisher, 2009; Vul & Pashler, 2008).

Of course, no one who reflects on personal experience of viewing a rectangle comes away thinking that his or her mind runs Blum's (1973) grassfire transform on it. So how is it that participants in the shape-tapping experiments were compelled to touch the medial axis? There is surely no inherent connection between tapping shapes and representing them: After all, people do not go around tapping shapes in daily life, and the purpose of skeletal shape representations is not to guide shape tapping. But such representations are thought to be deployed during literally every moment one's eyes are open, constructed by the visual system for every form and figure encountered. The reason touches of the shapes formed their medial axes, then, may be rather innocuous: Arbitrary and "unthinking" tasks with no obvious answer are precisely the contexts in which latent mental representations can subtly exert their influence. (Compare how spreading activation in memory is assessed via stem-completion priming, for example.) It has already been suggested, for example, that visual attention prioritizes shapes' symmetry axes (Barenholtz & Feldman, 2003; Harrison & Feldman, 2009; Kovacs et al., 1998; Kovacs & Julesz, 1994; Wang & Burbeck, 1998), so participants in our task may have picked the locations they did because those locations were attentionally enhanced.

We hope this surprisingly direct window onto otherwise-hidden visual processes will continue to bear fruit in illuminating how the visual system represents forms and figures, and that it will serve as a case study in the larger quest to understand the format and status of representations in the mind.

Author Contributions

C. Firestone and B. J. Scholl designed the study and wrote the manuscript. C. Firestone collected the data and conducted the analyses.

Acknowledgments

For helpful conversation or comments on earlier drafts, we thank Jacob Feldman, Pascal Mamassian, Manish Singh, and Tim Vickery.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Note

1. Modifying this analysis to consider the average distance to several nearest touches (e.g., the nearest 5 or 10 touches) did not qualitatively change this pattern of results (nor did it qualitatively change the result of any other analysis reported in this article).

References

Barenholtz, E., & Feldman, J. (2003). Visual comparisons within and between object parts: Evidence for a single-part superiority effect. *Vision Research*, *43*, 1655–1666.

Blum, H. (1973). Biological shape and visual science (Part I). *Journal of Theoretical Biology*, *38*, 205–287.

Burbeck, C. A., & Pizer, S. M. (1995). Object representation by cores: Identifying and representing primitive spatial regions. *Vision Research*, *35*, 1917–1930.

Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences, USA*, *103*, 18014–18019.

Harrison, S. J., & Feldman, J. (2009). The influence of shape and skeletal axis structure on texture perception. *Journal of Vision*, *9*(6), Article 13. Retrieved from http://www.journalofvision.org/content/9/6/13.full?sid=5789855f-4272-4e24-8884-ca5cbc7013e4

Hung, C.-C., Carlson, E. T., & Connor, C. E. (2012). Medial axis shape coding in macaque inferotemporal cortex. *Neuron*, *74*, 1099–1113.

Kimia, B. B. (2003). On the role of medial geometry in human vision. *Journal of Physiology-Paris*, *97*, 155–190.

Kovacs, I., Feher, A., & Julesz, B. (1998). Medial-point description of shape: A representation for action coding and its psychophysical correlates. *Vision Research*, *38*, 2323–2333.

Kovacs, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature*, *370*, 644–646.

Leyton, M. (1992). *Symmetry, causality, mind*. Cambridge, MA: MIT Press.

Liu, T. L., & Geiger, D. (1999). Approximate tree matching and shape similarity. In *Proceedings of the Seventh IEEE International Conference on Computer Vision* (Vol. 1, pp. 456–462). Los Alamitos, CA: IEEE Computer Society.

Melcher, D., & Kowler, E. (1999). Shapes, surfaces and saccades. *Vision Research*, *39*, 2929–2946.

Palmer, S. E., & Guidi, S. (2011). Mapping the perceptual structure of rectangles through goodness-of-fit ratings. *Perception*, *40*, 1428–1446.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*(10). Retrieved from http://www.frontiersin.org/Journal/10.3389/neuro.11.010.2008/full

Pizer, S. M., Siddiqi, K., Szekely, G., Damon, J. N., & Zucker, S. W. (2003). Multiscale medial loci and their properties. *International Journal of Computer Vision*, *55*, 155–179.

Psotka, J. (1978). Perceptual processes that may create stick figures and balance. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 101–111.

Shaked, D., & Bruckstein, A. M. (1998). Pruning medial axes. *Computer Vision and Image Understanding*, *69*, 156–169.

Siddiqi, K., & Pizer, S. (2008). *Medial representations: Mathematics, algorithms, and applications*. New York, NY: Springer.

Siddiqi, K., Shokoufandeh, A., Dickinson, S., & Zucker, S. (1999). Shock graphs and shape matching. *International Journal of Computer Vision*, *30*, 13–32.

van Tonder, G. J., Lyons, M. J., & Ejima, Y. (2002). Visual structure of a Japanese Zen garden. *Nature*, *419*, 359–360.

Vishwanath, D., & Kowler, E. (2003). Localization of shapes: Eye movements and perception compared. *Vision Research*, *43*, 1637–1653.

Vul, E., Hanus, D., & Kanwisher, N. (2009). Attention as inference: Selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General*, *138*, 546–560.

Vul, E., & Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, *19*, 645–647.

Wang, X., & Burbeck, C. A. (1998). Scaled medial axis representation: Evidence from position discrimination task. *Vision Research*, *38*, 1947–1959.

Wilder, J., Feldman, J., & Singh, M. (2011). Superordinate shape classification using natural shape statistics. *Cognition*, *119*, 325–340.

Zhu, S. C., & Yuille, A. L. (1996). FORMS: A flexible object recognition and modelling system. *International Journal of Computer Vision*, *20*, 187–212.