

# Where to Draw The Line?

Heping Sheng

Boston University

John Wilder

University of Toronto

Dirk B. Walther (✉ [bernhardt-walther@psych.utoronto.ca](mailto:bernhardt-walther@psych.utoronto.ca))

University of Toronto

---

## Research Article

**Keywords:** Drawings, Natural Scenes, Importance of Contours, Scene Perception

**Posted Date:** June 2nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-536728/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Where to draw the line?

Heping Sheng<sup>1</sup>, John Wilder<sup>2</sup>, Dirk B. Walther<sup>2\*</sup>

<sup>1</sup> School of Medicine, Boston University

<sup>2</sup> Department of Psychology, University of Toronto

\*Corresponding author: Dirk B. Walther

Email: [bernhardt-walther@psych.utoronto.ca](mailto:bernhardt-walther@psych.utoronto.ca)

**Keywords:** Drawings; Natural Scenes; Importance of Contours; Scene Perception

**Number of Words:** 4570

**Abstract** (200 words)

We often take people's ability to understand and produce line drawings for granted. But where should we draw lines, and why? We address fundamental principles that underlie efficient representations of complex information in line drawings. First, 58 participants with varying degree of artistic experience produced multiple drawings of a small set of scenes by tracing contours on a digital tablet. Second, 37 independent observers ranked the drawings by how representative they are of the original photograph. Overall, artists' drawings ranked higher than non-artists'. Matching contours between drawings of the same scene revealed that the most consistently drawn contours tend to be drawn earlier. We generated half-images with the most- versus least-consistently drawn contours by sorting contours by their consistency scores. Twenty five observers performed significantly better in a fast scene categorization task for the most compared to the least consistent half-images. The most consistent contours were longer and more likely to depict occlusion boundaries. Using psychophysics experiments and computational analysis, we confirmed quantitatively what makes certain contours in line drawings special: longer contours mark occlusion boundaries and aid rapid scene recognition. They allow artist and non-artists to convey important information starting from the first few strokes in their drawing process.

## Main Text

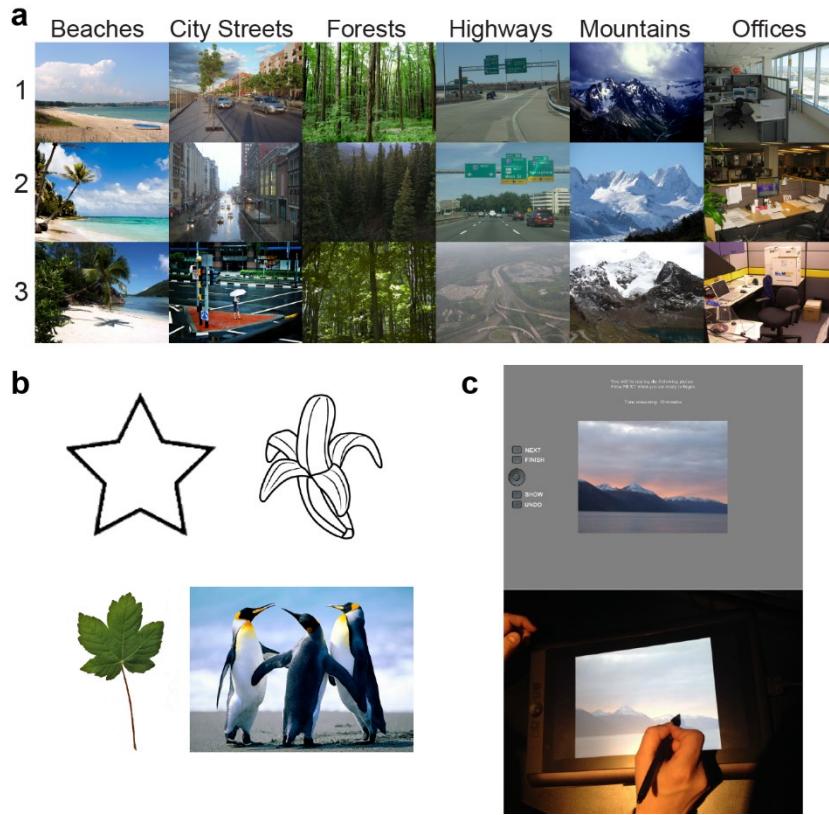
### Introduction

Humans have used line drawings to depict scenes from their lives for at least 45,500 years [1]. Line drawings can capture the essence of shapes and spatial relationships in complex scenes to an extent that makes them not only reasonable stand-ins for photorealistic images, but can emphasize particular aspects of the scene unencumbered by other, extraneous features that might otherwise obscure the clarity arising from a purely contour-driven depiction. For this reason, line drawings are used for technical drawings, such as architectural plans, designs for complex machinery, or charts for complex production processes. Even though line drawings lack texture and color, they nevertheless represent essential information, making them suitable for efficiently conveying visual information non-verbally, for instance in signs that request patrons to wear masks inside stores, instructions for the Heimlich maneuver, or step-by-step instructions for assembling a MALM chest of drawers from IKEA.

Line drawings of real-world scenes are recognized as quickly and accurately as photographs [2, 3]. Perception of line drawings is innate to human children [4–6] and does not rely on an acquired cultural framework [7, 8]. Even chimpanzees were found to understand line drawings without prior training [9].

In their first attempts to artistically reflect their world, children resort to line drawings [10]. Attempts at photorealistic painting tend to come at a later age. Neural representations of scenes, objects and faces elicited by line drawings were found to be equivalent to those elicited by the corresponding color photographs in visual cortex [11–13], although the equivalence for scenes appears to be restricted to the level of scene categories, not scene identity [14]. These similarities of line drawings to photorealistic depiction provide interesting insights into visual brain functions [15].

However, not all line drawings convey visual information equally well. The quality of a drawing, and thereby its usefulness to convey visual content, depends on the experience and on the particular choices of the artist. Lines are placed for different reasons, e.g., to indicate the boundaries of objects, structure within objects, occlusion of one object by another, texture elements, or the outlines of shadows cast by objects [16]. We here demonstrate that some lines are more important than others for conveying visual information, and we establish a tight relationship between the importance of contour lines, artistic expertise, and the ability of line drawings to convey scene content. In a series of three experiments, participants with varying degree of artistic experience were asked to trace the important contours in a set of photographs of complex real-world scenes (Experiment 1). Such a tightly controlled tracing experiment allows us to match contours between drawings of the same scene. Then, an independent group of participants rank-ordered the drawings belonging to the same photograph, allowing us to establish a measure of contour importance (Experiment 2). Finally, we test the effect of contour importance for scene categorization (Experiment 3).



**Figure 1.** Stimuli and setup for Experiment 1. (a) Photographs used for generating line drawings: the top 3 most representative photographs from each of six scene categories, chosen from a larger set of photographs. (b) Practice stimuli to familiarize participants with the tracing task and program interface. (c) Experiment setup with a pen display graphics tablet. The computer screen displayed instructions and the original 800x600 pixels photograph as reference during the trial (left). The graphics tablet displayed a 1440 x 1080 pixels copy of the image with a semi-transparent overlay for tracing (right). The Next, Finish, Show, and Undo functions corresponding to each button on the side of the graphics tablet were displayed during the entire experiment.

## Results

### *Experiment 1: Tracing scene photographs*

We sought to determine how consistent individuals are at drawing lines as well as effects of artistic expertise on tracing photographs. To this end, we performed a controlled drawing experiment, in which we asked participants with varying degrees of artistic experience to trace the important contours in a set of 18 photographs of real-world scenes (Figure 1a). We chose a tracing task rather than free drawing or copying the scenes from the photographs so that we could match the contours in the drawings by different participants based on their spatial location. The task still gives participants freedom to decide which lines they deem important enough to include in their drawings. We also piloted free-hand copying of scenes and drawing from memory. However, these tasks are significantly more difficult and especially daunting for participants without artistic training. Participants were even overwhelmed when presented

with the task of tracing a complex scene without preparation. We therefore eased them into the experimental procedure by using successively more difficult tracings, starting with a simple geometric shape (Figure 1b).

Fourteen students majoring in visual arts or architecture from the University of Toronto and the Ontario College of Art and Design University (artists) were asked to trace the important contours of photographs of natural scenes on a graphics tablet (Figure 1c). One artist was excluded from the analysis, because she obviously did not make an effort to produce recognizable drawings. The remaining 13 artists (12 women; age 18-48, mean age 25; 3 left-handed) reported to have 4 to 32 years of arts experience (median 10), had taken 0.5 to 20 full-year arts courses (median 4.5), and draw for 1 to 52 hours per week (median 5). We asked an additional 44 undergraduate students majoring in psychology at the University of Toronto (non-artists) to draw contours of the scenes. We excluded two non-artists from the analysis – one for adding handwritten words to the drawings, and one for exceedingly sloppy work. The remaining 42 non-artists (26 female, age 17-24 years, mean age 19 years; 2 left-handed) reported up to 12 years of arts experience (median 1) but did not report having taken any arts courses or regularly practicing drawing.

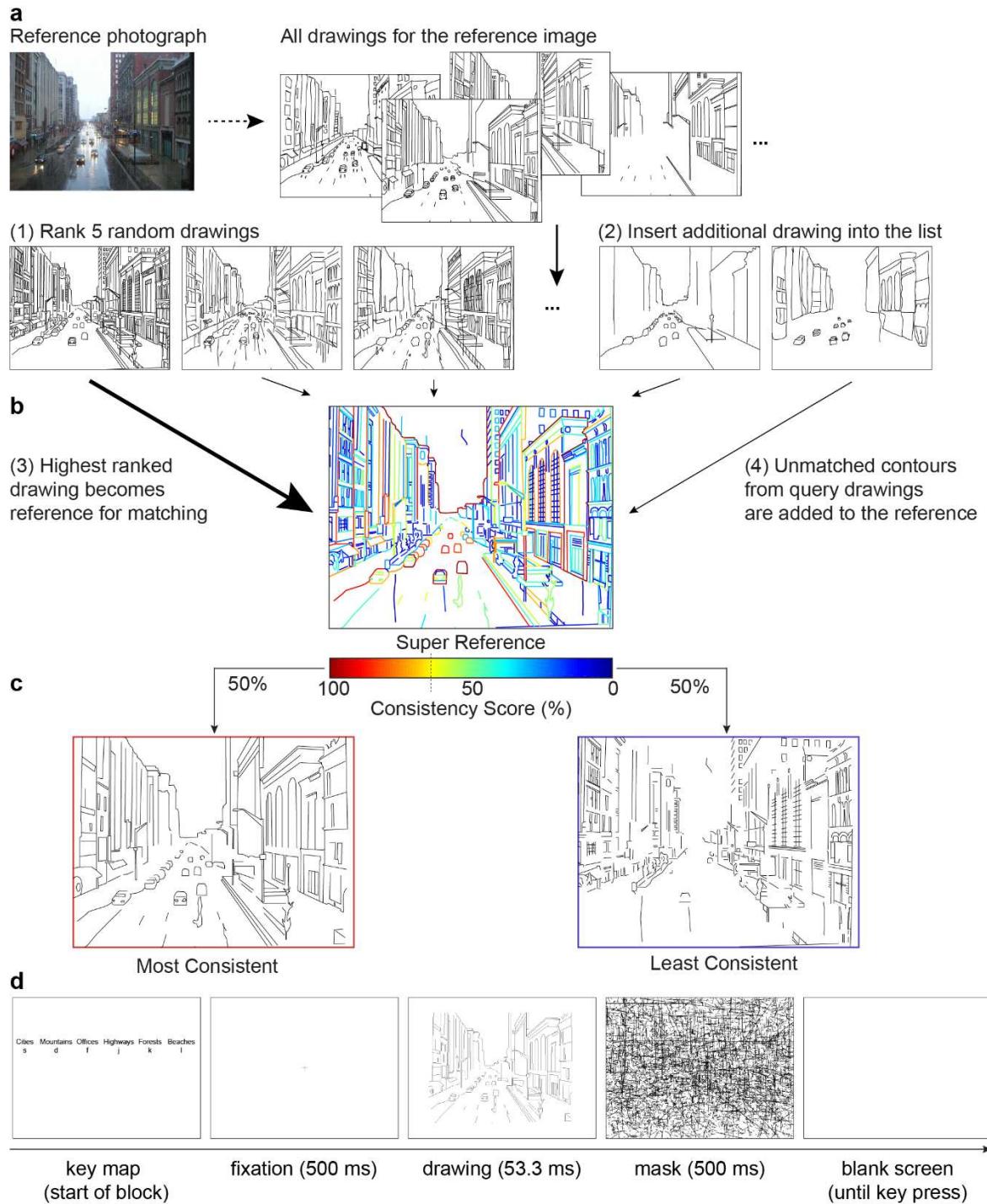
As expected, artists had more years of self-reported art experience, with a median of 10 years compared to 1 year for controls ( $t(12) = 5.75, p < 0.001$ , two-sample t test). The artist group was older (25.23 versus 19.24 years), more female predominant (92% vs 62%), and less right hand dominant (77% vs 95%). We collected a total of 194 line drawings from artists and 159 from non-artists. We added 18 drawings commissioned from trained artists [13] for a grand total of 371 line drawings, with 15 to 24 (median 21) drawings for each photograph.

### *Experiment 2: Ranking line drawings*

The quality of the line drawings collected in Experiment 1 varied considerably. For an impartial assessment, we performed a ranking experiment with an independent group of observers (Figure 2a). Each set of line drawings was ranked by 5 to 12 participants (median 9). Participants strongly agreed on how much each line drawing represents the original photograph, as evidenced by the significant coefficients of concordance that ranged between 0.57 and 0.92 (Table S1).

Furthermore, we established the match of contours between drawings of the same scene. We constructed a super-reference as the superset of unique contours from all drawings of the same scene. The match between contours allowed us to count how frequently a particular contour in the super-references was drawn across participants in Experiment 1. We used this drawing frequency as a proxy for the consistency of contour lines (Figure 2b). All 18 super-references are shown in Figure S1, with color of the contours coding for their consistency scores.

We performed a multiple linear regression analysis to predict average normalized rank scores ( $S_R$ ) of individual line drawings based on artistic experience ( $X$ ) of the author of the drawing (measured in number of years of artistic experience) and based on the total number of pixels ( $N_P$ ) in the drawing as a covariate. We excluded four drawings by one outlier artist with 32 years of experience from the analysis. The range of artistic experience without this artist was 0 – 17 years. We also did not include the drawings originally commissioned for [13], since we had no information about the authors of these drawings.



**Figure 2.** Procedures for Experiments 2 and 3. (a) Ranking of line drawings was performed by rearranging images printed on cardstock on a large table. (b) Constructing a super-reference by matching contours in the individual drawings. (c) Construction of half drawings by splitting up contours according to consistency scores. (d) Experimental procedure for the six-alternative forced-choice scene categorization experiment.

**Table 1.** Regression table for the regression in Equation 1.

	coefficient	SE	t-stat	p-value
Intercept	-3.214	3.902	-0.824	0.411
Artistic Experience	2.092	0.626	3.344	$9.17 \cdot 10^{-4}$
Number of Pixels	2.633	0.246	10.719	$< 2 \cdot 10^{-16}$
Experience*Pixels	-0.045	0.037	-1.228	0.220

We found a significant regression equation ( $F(3, 345) = 104.8, p < 0.001$ ) with an adjusted  $R^2$  of 0.472 (Figure S2). The resulting regression equation is:

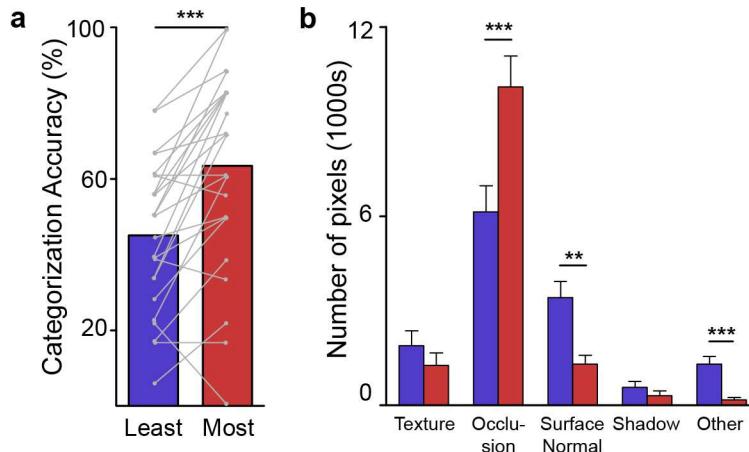
$$\hat{S}_R = -3.214 + 2.092 \cdot X + 2.633 \cdot N_P - 0.045 \cdot X \cdot N_P \quad (1)$$

where artistic experience  $X$  is expressed in years and  $N_P$  in 1000s of pixels. Artistic experience and total number of pixels were significant predictors of rank with positive coefficients (Table 1). That is, more detailed drawings (with more pixels) and drawings by participants with more experience were ranked more highly. Importantly, the two factors contributed to the ranking of the drawings independently, as there was no significant interaction. That is, while trained artists tend to produce more detailed drawings, the advantage of artistic training is not explained by the amount of detail alone, since both factors contribute to the rank of the drawings independently.

One way to equalize the amount of detail would be to modify the experiment to limit the amount of “ink” (black pixels) available to participants when they trace the scenes. However, such an approach is likely to introduce additional confounds regarding the economizing of ink as a scarce resource, such as anxiety about using the ink wisely, especially for untrained participants. Furthermore, the amount of available ink would need to vary by the complexity of the underlying scene, making comparisons between scenes challenging. We therefore opted for the current approach of not limiting ink, which leads to differences in the amount of detail between participants.

If individuals show such high agreement on a certain set of contours, do they also draw these lines earlier in the trial? To answer this question, we divided the time when a line was drawn within a trial by the total duration of the trial, resulting in relative timing in the range between zero and one for each of the 353 drawings produced in Experiment 1. Since we did not have timing information for the 18 drawings from [13], we excluded these drawings from the analysis. Consistency scores of contours in the super-reference drawings were mapped back to the matching contours in the individual drawings.

We fitted a linear mixed-effects model to explain the relationship of consistency scores of individual contours (58480 contours from 353 drawings) with relative timing (fixed effect), with drawing identity as a random-effect covariate. We found a highly significant negative effect of timing ( $\beta = -0.206$ ; 95% confidence interval = [-0.213, -0.199];  $t(58478) = -55.88; p < 0.001$ ), with a smaller random effect of image identity ( $\beta = 0.101$ ; CI = [0.093, 0.109]). This result shows that the most consistent lines tend to be drawn earlier, suggesting that both the consistency and the timing of drawings reflect their importance for perception. We test this hypothesis explicitly in Experiment 3.



**Figure 3.** Results of Experiment 3. (a) Categorization accuracy for least and most consistent half drawings. (b) Number of pixels in particular types of contours in least (blue) and most (red) consistent half drawings. \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

#### *Experiment 3: Effect of line consistency for scene perception*

In Experiment 2 we computed the consistency score by matching the drawings of the same photograph by several participants to a super-reference drawing. Does a high consistency score for a particular contour mean that the contour is more important for representing scene content? We address this question with a fast scene categorization experiment with stimuli designed to maximize the difference in consistency scores.

Since drawn contours represent what participants deemed to be important and salient edges in the scene, we hypothesized that the frequency at which a contour is drawn across participants is a good measure of its relative importance for perceiving the scene. To test this hypothesis, we used the consistency scores computed for the super-references in Experiment 2 to construct half-drawings that contained either the most or the least consistent half of the contours (Figures 2c; S3). We used these half-drawings as stimuli for a six-alternative forced-choice (6AFC) scene categorization experiment (Figure 2d). We hypothesized that the most consistent half-drawings would enable more accurate recognition compared to the least consistent half drawings.

Furthermore, we sought to determine what distinguishes more important contour lines from less important ones. Contours in a drawing can have different physical causes in the real scene. For instance, they can represent depth discontinuities, such as occlusion boundaries, relate to changes in surface curvature, or be part of textures. We here investigate whether any of these physical roles are more important for conveying scene information than others in drawings of complex real-world scenes.

Participants categorized the half-drawings containing the contours with the lowest consistency scores with 43% accuracy (chance: 16.7 %) and the half-drawings with the high-consistency contours with 63% accuracy (Figure 3a). The difference was highly significant ( $t(14) = 4.21$ ;  $p < 0.001$ ).

What makes the most consistently drawn contours so much better at conveying scene information? We annotated all contours in the super references according to their physical cause and analyzed the types separately for the least and the most consistent half drawings (Figure 3b). A 2x5 analysis of variance

(ANOVA) with most/least consistent and contour type as factors showed a significant main effect for contour type ( $F(4,170) = 79.59, p < 0.001$ ) but not for most/least consistent half drawing ( $F(1,170) = 0.014, p = 0.906$ ), as expected since the half drawings are by design equated in the total number of pixels. Importantly, we found a significant interaction ( $F(4,170) = 10.98; p < 0.001$ ). The most consistent half drawings contained significantly more pixels belonging to occlusion boundaries, typically object boundaries, than the least consistent half drawings ( $t(17) = 8.346, p < 0.001$ ; Bonferroni-corrected for multiple comparisons). This was balanced by more pixels for the least than the most consistent half drawings belonging to contours associated with surface normals ( $t(17) = -4.28, p = 0.00254$ ) and contours that could not be clearly assigned ( $t(17) = -4.78, p < 0.001$ ). There was no difference in the number of pixels associated with texture boundaries ( $t(17) = -1.94, p = 0.348$ ) or cast shadows ( $t(17) = -1.65, p = 0.588$ ).

## Discussion

In the three experiments presented here we have established that (i) line drawings are rated as more representative of the depicted scene when they are drawn by experienced artists; (ii) the most consistently drawn contours are drawn earlier; (iii) drawings containing the most consistently drawn contours are more recognizable than drawings with the least consistently drawn contours; and (iv) the most consistent contours are more likely than inconsistent contours to represent occlusion boundaries, that is, the boundaries of objects.

While these results may seem unsurprising in hindsight, this is the first time, to our knowledge, that the ability of humans of various levels of artistic expertise to convey essential information in line drawings has been quantitatively measured for complex real-world scenes. Specifically, we used artistic expertise in the participant populations as a control variable to confirm that artistic experience does indeed result in drawings that are objectively more recognizable. To this end, we established an algorithmic measure of the importance of contours by determining how consistently contours were drawn across all participants. More consistently drawn contours tended to be drawn earlier in the drawing process. Importantly, the fast scene categorization task in Experiment 3 established the perceptual advantage in an unbiased, objective way, since participants in that study were not aware of the consistency condition of the stimulus drawings. Finally, we found that this perceptual advantage was tied to occlusion boundaries, which represent the shapes of objects in the scenes (Figure 3b).

Human proficiency at perceiving objects in complex scenes has been previously linked to edges created from surface normals and depth boundaries in a study with synthetic scenes [16]. Our analysis of the roles of contours in Experiment 3 confirms these findings in a real-world scene setting. These findings are consistent with the important role of contour junctions for the perception of objects and scenes, as contour junctions serve as a low-level cue to spatial relations such as occlusion [17–19]. Moreover, we show how artists prioritize contours that lead to drawings which are more representative of the depicted scene (Equation 1) and that lead to better perception of scene gist (Figure 3a). This finding is likely related to the technique of “blocking-in” – a coarse, block-like outline of the proportions of figures and objects in the initial phase of drawing [20]. This technique may lead trained artists to initially prioritize contours that convey global shape over contours that convey finer details.

In our work we chose a tracing task in order to obtain multiple drawings for each of the complex scenes, for which we would be able to establish the correspondence of the contours and, by extension, their

importance as a measure of relative frequency. While this approach affords us the experimental control required for our study, it also has the drawback of limiting the artistic choices by the participants. Previous work has shown that when asked to produce a line drawing of a recently viewed scene from memory, non-artists draw many of the object boundaries, and in roughly the correct location [21]. This suggests that humans intuitively know which lines are important to convey the meaning of a scene, but that artistic training improves upon this ability. Recent computer vision work has demonstrated that non-artists can be trained to prioritize the most important lines, and draw them earlier than unimportant lines [22]. Furthermore, even artificial sketch generators, when trained to create a sketch to convey the essence of an image with as few strokes as possible learn to first draw lines that have the most power to convey essential content [22]. In fact, there have recently been a number of artificial neural networks trained to generate sketches that are as easily recognizable as those generated by a human [23, 24]. We here show that drawings that take advantage of the visual system’s mechanisms for understanding scenes will be more easily interpreted [25].

It is important to point out that we are here not addressing any issues of artistic style, artistic expression or their relationship to perceived aesthetic value. In particular, we make no claims of labeling the line drawings created in this highly controlled study as “artwork.” Visual artistic expression, although sometimes concerned with faithful representation of the real world, involves many more aspects, such as composition, emotional content, and frequently metaphorical allusions that transcend the figurative content of the physical artwork. Nevertheless, our findings may help to illuminate what attributes make artwork recognizable, often despite extreme distortions or extreme simplification – for example, gesture drawings or Cubist paintings.

To conclude, we have presented a set of controlled experiments on the production and perception of line drawings for conveying the content of complex real-world scenes. We found that contours drawn most consistently across individuals are most effective at conveying scene content, that contours drawn earlier in the drawing process show higher consistency, and that trained artists are more likely to draw consistent contours. More consistent contours are more likely than less consistent contours to convey occlusion boundaries, which signal the shape of objects in a scene as well as their spatial relationships.

## Materials and Methods

### *Experiment 1*

We recruited two groups of participants: 14 students majoring in visual arts or architecture from the University of Toronto and the Ontario College of Art and Design University (artists) participated in a two-hour drawing experiment for monetary compensation. In addition, 44 undergraduate students majoring in psychology at the University of Toronto (non-artists) participated in the study for one hour for course credit. The sample size of the artist group was determined by the number of artists who responded to our call for participation. We chose the size of the control sample to approximately equate the total number of drawings produced by the two groups.

We asked all participants to indicate how many years of experience in producing visual arts they had, how many full-year arts courses they had taken, and how much time they typically spend drawing per week. All participants reported normal or corrected-to-normal vision and provided written informed consent. The experiment was approved by the Research Ethics Board of the University of Toronto (Protocol #30999) and followed the guidelines set out in the Declaration of Helsinki.

The 18 images used in this experiment were chosen from a set of photographs of real-world scenes [13]. The images have been rated by an independent group of participants on how well they represent one of six scene categories: beaches, city streets, forests, highways, mountains, and offices. Since knowledge of category membership may aid people in constructing a clearer representation of the scene [26], the top three exemplar photographs from each scene category were chosen for a total of 18 stimulus images (Figure 1A). Four images distinct from the experimental categories were used for practice trials: a regular concave decagon (five-pointed star), a line drawing of a banana, a photograph of a leaf in isolation, and a photograph of penguins with background (Figure 1B).

Participants were asked to draw outlines of the original images onto a semitransparent overlay using a Wacom Cintiq 13HD Interactive Pen Display graphics tablet (Figure 1C). We recorded spatial coordinates and timing of the strokes for all lines in the final drawings (see Supplemental Online Materials for details). Participants who made unrecognizable drawings or drawings with excessive extraneous features were excluded from further analysis. All drawings were scaled to 800x600 pixels, and line segments of zero length removed.

### *Experiment 2*

We recruited a separate group of 37 undergraduate students of psychology from the University of Toronto. Twenty-five participants spent one hour on the experiment in lieu of partial course credit. Twelve participants volunteered their time and spent less than one hour, performing fewer ranking trials. This sample size was estimated to produce approximately 10 rankings for each of the 18 original images. Participants were unfamiliar with the photographs and line drawings in the study. All participants reported normal or corrected-to-normal vision and provided written informed consent. The experiment was approved by the Research Ethics Board of the University of Toronto (Protocol #30999) and followed the guidelines set out in the Declaration of Helsinki.

We printed all 371 line drawings from Experiment 1 on cardstock (15 x 10 cm) and grouped them according to the identity of the original photographs. Each printed drawing was labeled on the back with a unique identifier. We also printed the 18 original color photographs for reference. Participants were asked to sort each set of drawings from most to least representative of the photo by physically arranging the printed drawings on a table and writing down the identifiers in order. To reduce the difficulty of comparing up to 24 drawings at once, raters were instructed to first rank 5 randomly selected drawings, and then insert each additional one into its appropriate location in the list (Figure 2A). Participants proceeded through a randomly selected subset of image identities at their own pace, finishing between 1 and 7 sets (median 6).

After collecting all rankings, each drawing was assigned a normalized rank score, with 100% indicating high rank and 0% low rank with respect to all drawings produced for a particular photo. To examine the consistency among raters, we computed Kendall's coefficient of concordance for all rankings of drawings for each image, and performed significance testing using chi-squared statistics. Trials that did not result in a complete ranking of a set of drawings were excluded from the analysis. For further analysis, we averaged the normalized rank scores for each drawing across raters. The most representative drawing for each image was chosen as the reference for subsequent contour matching between drawings of the same original photograph.

We examined the factors that may have contributed to the ranking with a multiple linear regression analysis. We sought to predict ranking of a particular line drawing based on the artistic experience of its author and the total number of pixels in the drawing, as it is plausible that the amount of detail in a drawing also affects perceived representativeness.

Using a matching algorithm that accounted for proximity of line segments as well as consistency of matches across longer contours, we established which contours in one drawing match which contours of another drawing of the same scene. From these matches we constructed super-references that contained the superset of contours from all drawings of the same scene (Figure 2B). We established a consistency score for each contour in the super-reference of a scene by counting the number of matching contours across all drawings, divided by the total number of drawings for the scene. See Supplemental Online Materials for details on the contour matching and the construction of the super-reference.

### *Experiment 3*

Twenty-five undergraduate students of Psychology (ages 17-28, mean 18.9; 16 female) at the University of Toronto participated in the study for partial course credit. This experiment was performed as an add-on for another, similarly structured main experiment. The sample size was determined based on the requirements of the main experiment. All participants reported normal or corrected-to-normal vision and provided written informed consent. The experiment was approved by the Research Ethics Board of the University of Toronto (Protocol #30999) and followed the guidelines set out in the Declaration of Helsinki.

In order to examine the perceptual characteristics of the most-consistently and least-consistently drawn contours, we divided each of the 18 super-references into a high- and a low-consistency half-drawing according to their consistency scores as computed in Experiment 2. The two drawings did not share any contours, and each contained approximately 50% of the total pixels in the original super-reference (Figure 2C). As a result, we obtained 36 half-line drawings, 18 with the most consistent contours (Figure S3A) and 18 with the least consistent contours (Figure S3B).

Participants performed a fast scene categorization experiment with an image presentation time of 53 ms, followed by a perceptual mask. They were asked to press one of six randomly assigned keys to indicate which scene categories they saw. For details more details on equipment and procedures see Supplemental Online Materials.

To analyze contour types, the first author manually labelled each contour in the super-reference drawings by sequentially overlaying them over the original photograph. She classified contours into four different edge types according to their physical cause [16]: texture/albedo edges (change in reflectance across smooth surface), occlusion/depth boundaries (boundaries of objects), surface normal discontinuities (intersecting surfaces or creases), and shadow edges (boundary of cast shadows). For cases where a contour included more than one type of origin, the type that corresponds to the longer portion was chosen. Any contours that could not be assigned clearly to one of these types were labeled as “other”. Then, we totaled the number of pixels in contours belonging to each type within the most and least consistent half line drawings, and performed significance testing with a fixed-effects two-way ANOVA followed by Bonferroni-corrected paired t tests.

All data, statistical analyses and images of the drawings are available at: <https://osf.io/x9uj5/>

## Acknowledgments

We thank Profs. Natalie Waldburger and Amy Swartz for insightful discussions and for help with recruiting participants from OCAD University. This work was supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2020-04097) and the Social Sciences and Humanities Research Council of Canada (430-2017-01189).

## References

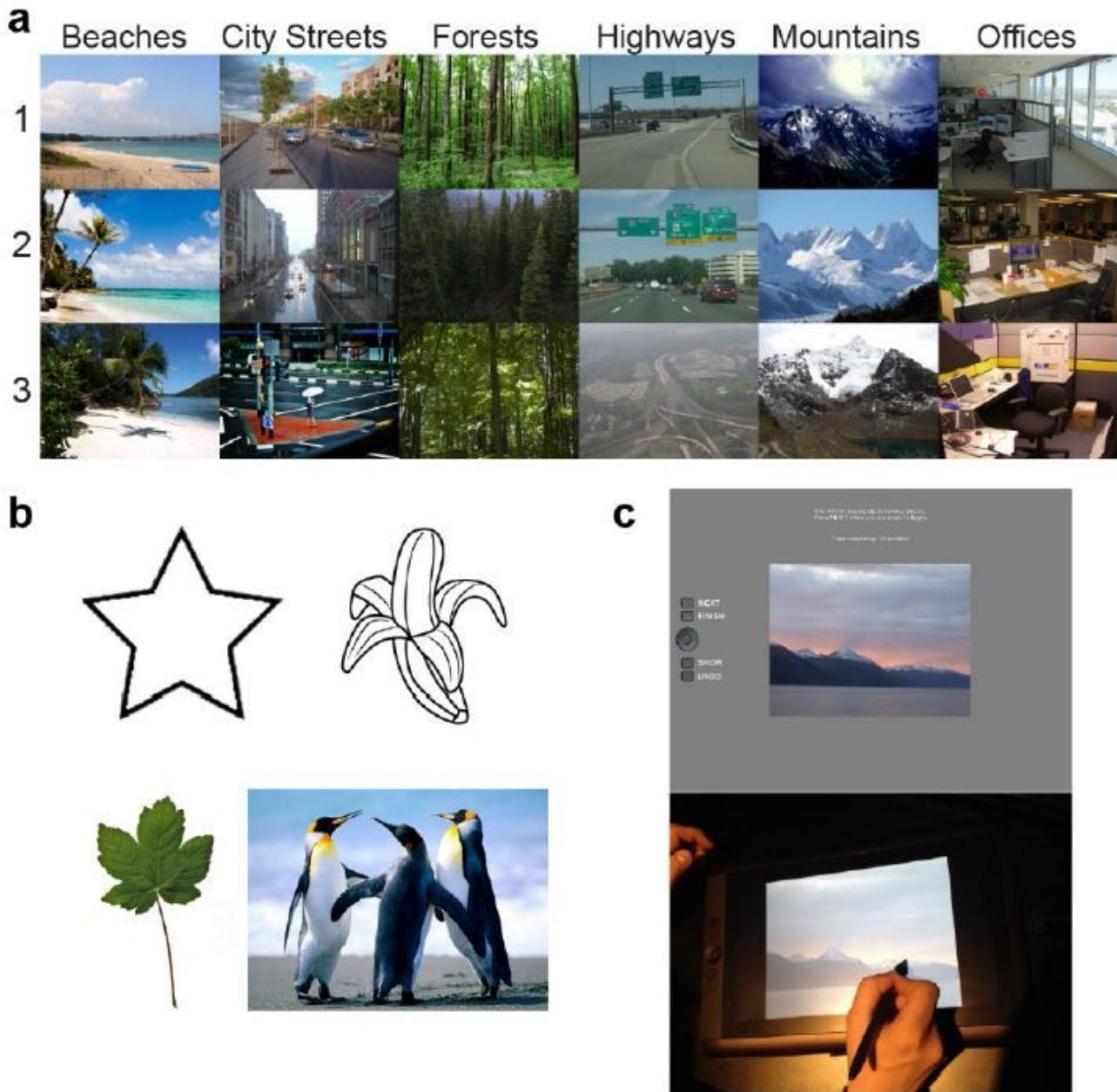
1. Brumm A, Oktaviana AA, Burhan B, Hakim B, Lebe R, Zhao J, Sulistyarto PH, Ririmasse M, Adhityatama S, Sumantri I, Aubert M (2021) Oldest cave art found in Sulawesi. *Science Advances* 7: <https://doi.org/10.1126/sciadv.abd4648>
2. Biederman I, Ju G (1988) Surface versus edge-based determinants of visual recognition. *Cognitive psychology* 20:38–64
3. Potter MC, Levy EI (1969) Recognition memory for rapid sequences of pictures. *Journal of experimental psychology* 81:10–15
4. Hochberg J, Brooks V (1962) Pictorial recognition as an unlearned ability: A study of one child's performance. *The American Journal of Psychology* 75:624–628
5. Jahoda G, Deręgowski J, Ampene E, Williams N (1977) Pictorial recognition as an unlearned ability. In: *The child's representation of the world*. Springer, pp 203–217
6. Yonas A, Arterberry ME (1994) Infants perceive spatial structure specified by line junctions. *Perception* 23:1427–1435
7. Kennedy JM, Ross AS (1975) Outline picture perception by the Songe of Papua. *Perception* 4:391–406
8. Kennedy, John M (1974) *A Psychology of Picture Perception: Images and Information*. Jossey-Bass Publishers
9. Itakura S (1994) Differentiated responses to different human conditions by chimpanzees. *Perceptual and Motor Skills* 79:1288–1290
10. Goodnow J (1980) *Children's Drawing (The Developing Child)*. HarperCollins Publishers
11. Ishai A, Ungerleider LG, Haxby JV (2000) Distributed neural systems for the generation of visual images. *Neuron* 28:979–990
12. Spiridon M, Kanwisher N (2002) How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron* 35:1157–1165
13. Walther DB, Chai B, Caddigan E, Beck DM, Fei-Fei L (2011) Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proceedings of the National Academy of Sciences* 108:9661–9666 . <https://doi.org/10.1073/pnas.1015666108>
14. O'Connell TP, Sederberg PB, Walther DB (2018) Representational differences between line drawings and photographs of natural scenes: A dissociation between multi-voxel pattern analysis and repetition suppression. *Neuropsychologia* 117:513–519 . <https://doi.org/10.1016/j.neuropsychologia.2018.06.013>

15. Sayim B, Cavanagh P (2011) What line drawings reveal about the visual brain. *Frontiers in human neuroscience* 5:118
16. Bansal A, Kowdle A, Parikh D, Gallagher A, Zitnick L (2013) Which edges matter? In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. pp 578–585
17. Biederman I (1987) Recognition-by-components: A theory of human image understanding. *Psychological Review* 94:115–147
18. Choo H, Walther DB (2016) Contour junctions underlie neural representations of scene categories in high-level human visual cortex. *NeuroImage* 135:32–44 . <https://doi.org/10.1016/j.neuroimage.2016.04.021>
19. Walther DB, Shen D (2014) Nonaccidental Properties Underlie Human Categorization of Complex Natural Scenes. *Psychological Science*. <https://doi.org/10.1177/0956797613512662>
20. Sanmiguel D (2003) *Art of Drawing: The Complete Course*, 2nd edition. Sterling, New York, NY
21. Bainbridge WA, Hall EH, Baker CI (2019) Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature communications* 10:1–13
22. Bhunia AK, Das A, Riaz Muhammad U, Yang Y, Hospedales TM, Xiang T, Gryaditskaya Y, Song Y-Z (2020) Pixelor: A Competitive Sketching AI Agent. So you think you can beat me? In: *SIGGRAPH Asia*
23. Das A, Yang Y, Hospedales T, Xiang T (2020) *BézierSketch: A generative model for scalable vector sketches*. Springer, pp 632–647
24. Gao C, Liu Q, Xu Q, Wang L, Liu J, Zou C (2020) SketchyCOCO: image generation from freehand scene sketches. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp 5174–5183
25. Hertzmann A (2020) Why do line drawings work? a realism hypothesis. *Perception* 49:439–451
26. Torralbo A, Walther DB, Chai B, Caddigan E, Fei-Fei L, Beck DM (2013) Good exemplars of natural scene categories elicit clearer patterns than bad exemplars but not greater BOLD activity. *PLoS One* 8:e58594 . <https://dx.doi.org/10.1371%2Fjournal.pone.0058594>

**Author Contributions:** HS and DBW designed the experiments; HS implemented the experiments; HS and JW collected the data; HS, JW, and DBW analyzed the data; HS and DBW wrote the manuscript; HS, JW and DBW edited the manuscript. The final version of the manuscript was approved by all authors.

**Competing Interest Statement:** The authors declare no competing interests.

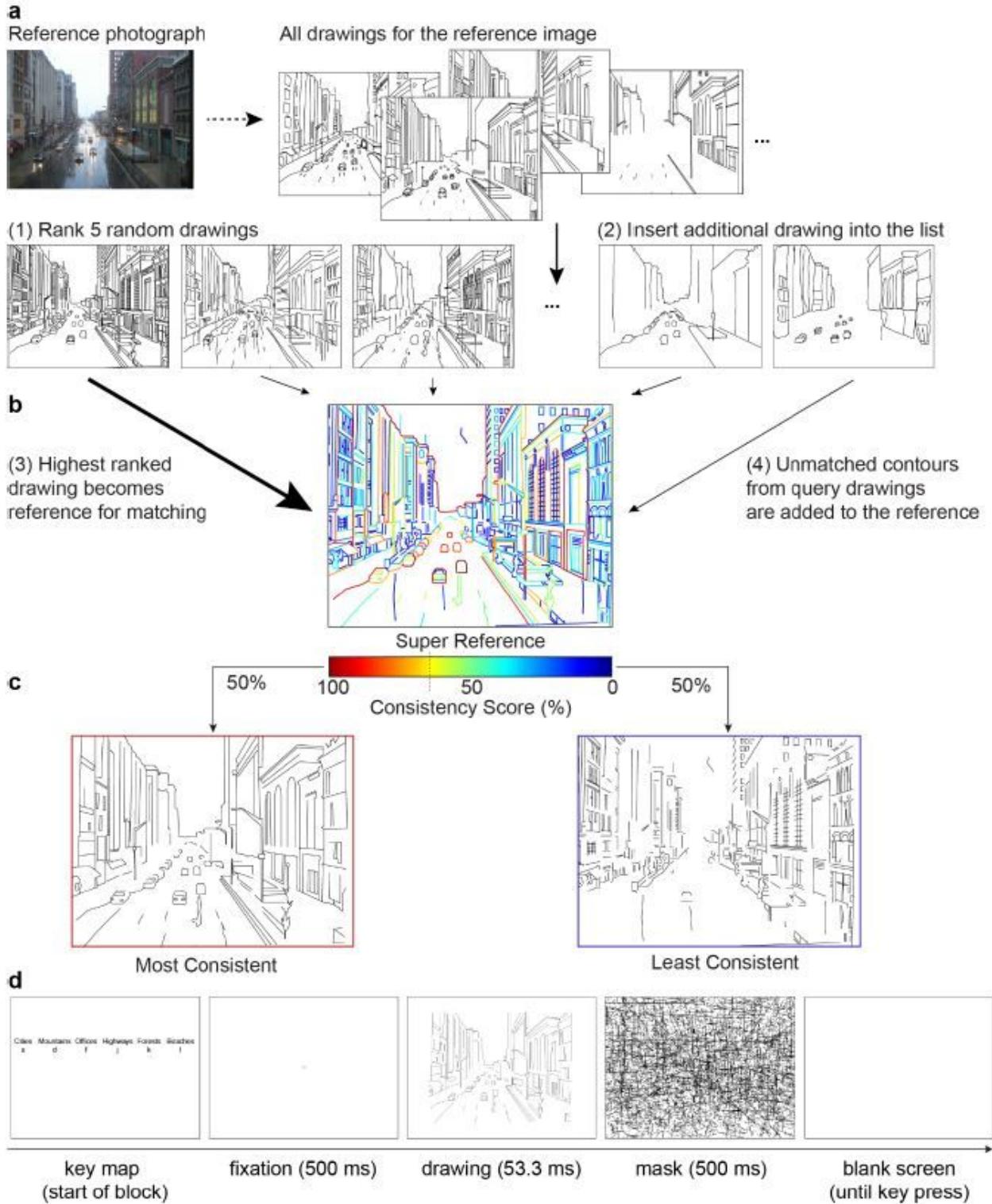
# Figures



**Figure 1**

Stimuli and setup for Experiment 1. (a) Photographs used for generating line drawings: the top 3 most representative photographs from each of six scene categories, chosen from a larger set of photographs. (b) Practice stimuli to familiarize participants with the tracing task and program interface. (c) Experiment setup with a pen display graphics tablet. The computer screen displayed instructions and the original 800x600 pixels photograph as reference during the trial (left). The graphics tablet displayed a 1440 x 1080 pixels copy of the image with a semi-transparent overlay for tracing (right). The Next, Finish, Show,

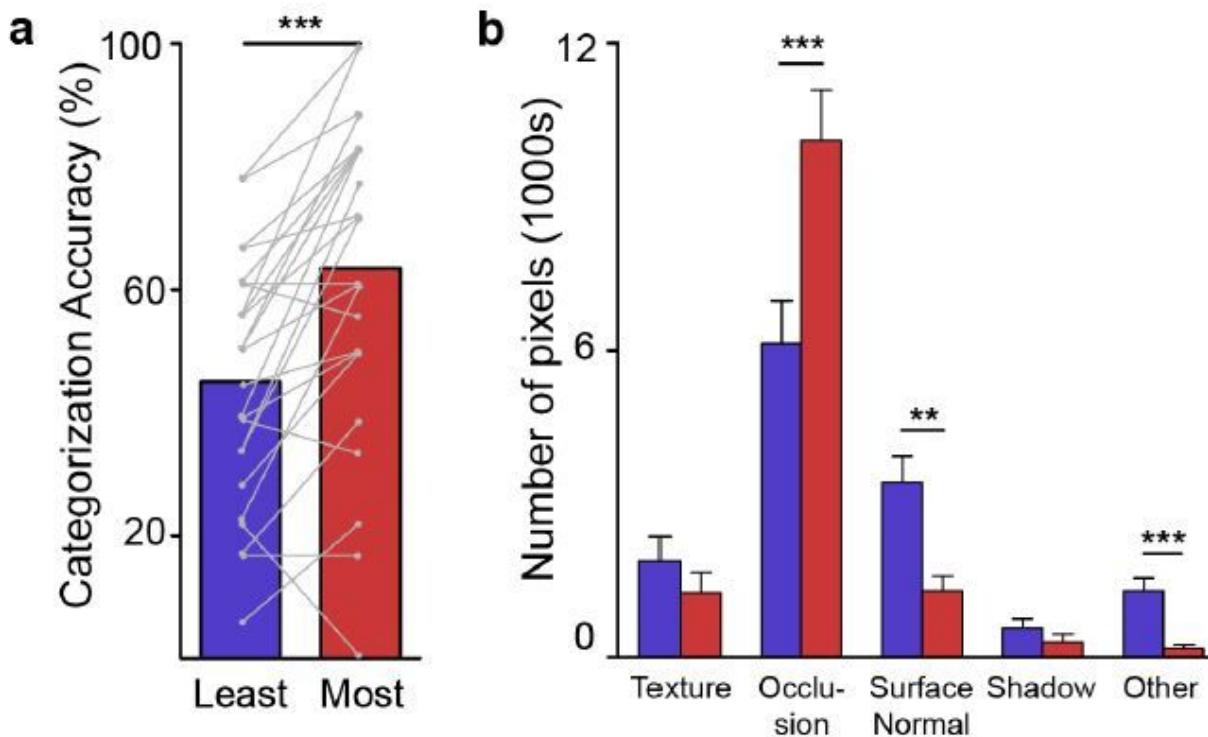
and Undo functions corresponding to each button on the side of the graphics tablet were displayed during the entire experiment.



**Figure 2**

Procedures for Experiments 2 and 3. (a) Ranking of line drawings was performed by re-arranging images printed on cardstock on a large table. (b) Constructing a super-reference by matching contours in the

individual drawings. (c) Construction of half drawings by splitting up contours according to consistency scores. (d) Experimental procedure for the six-alternative forced-choice scene categorization experiment.



**Figure 3**

Results of Experiment 3. (a) Categorization accuracy for least and most consistent half drawings. (b) Number of pixels in particular types of contours in least (blue) and most (red) consistent half drawings.  
\*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [WhereToDrawTheLineSciRepSuppl.pdf](#)