




Cognitive Science 45 (2021) e12933

© 2021 Cognitive Science Society, Inc All rights reserved.

ISSN: 1551-6709 online

DOI: 10.1111/cogs.12933

Curious Objects: How Visual Complexity Guides Attention and Engagement

Zekun Sun, Chaz Firestone *Department of Psychological & Brain Sciences, Johns Hopkins University*

Received 6 November 2019; received in revised form 10 December 2020; accepted 15 December 2020

Abstract

Some things look more complex than others. For example, a crenulate and richly organized leaf may seem more complex than a plain stone. What is the nature of this experience—and why do we have it in the first place? Here, we explore how object complexity serves as an efficiently extracted visual signal that the object merits further exploration. We algorithmically generated a library of geometric shapes and determined their complexity by computing the cumulative surprisal of their internal skeletons—essentially quantifying the “amount of information” within each shape—and then used this approach to ask new questions about the perception of complexity. Experiments 1–3 asked what kind of mental process extracts visual complexity: a slow, deliberate, reflective process (as when we decide that an object is expensive or popular) or a fast, effortless, and automatic process (as when we see that an object is big or blue)? We placed simple and complex objects in visual search arrays and discovered that complex objects were easier to find among simple distractors than simple objects are among complex distractors—a classic search asymmetry indicating that complexity is prioritized in visual processing. Next, we explored the function of complexity: Why do we represent object complexity in the first place? Experiments 4–5 asked subjects to study serially presented objects in a self-paced manner (for a later memory test); subjects dwelled longer on complex objects than simple objects—even when object shape was completely task-irrelevant—suggesting a connection between visual complexity and exploratory engagement. Finally, Experiment 6 connected these implicit measures of complexity to explicit judgments. Collectively, these findings suggest that visual complexity is extracted efficiently and automatically, and even arouses a kind of “perceptual curiosity” about objects that encourages subsequent attentional engagement.

Keywords: Complexity; Shape perception; Curiosity; Visual search; Information theory

*Correspondence should be sent to Zekun Sun or Chaz Firestone, Department of Psychological & Brain Sciences, Johns Hopkins University, 3400 N Charles St., Baltimore, MD 21218. E-mail: zekun@jhu.edu or chaz@jhu.edu

1. Introduction

Objects, events, and circumstances in the world frequently strike us as being *simple* or *complex*. For example, we can appreciate the complexity of a story's twisting plot, a city's interconnected highway system, or even a scientific theory's postulates and assumptions. In evaluating the complexity of these various systems and situations, we may attribute to them a kind of richness or density of information; for example, a story's plot may seem complex to the extent that it is difficult to summarize or predict, or when a single storyline branches into multiple subplots.

However, beyond these very abstract notions of complexity, which may require sustained reflection and deliberation to arise, the world can also *look* complex in a far more immediate, automatic, and even perceptual way. For example, consider the pairs of images in Fig. 1A. Over and above their particular low-level features (e.g., their shapes, orientations, or colors), the objects on the right evoke an impression of being more complex than the objects on the left. This impression is ubiquitous and robust, in that it can span a wide range of categories—here including both human-made artifacts and natural objects—as well as a variety of features and scales (including, e.g., both a massive tree and a tiny flower). But perhaps even more telling is that the experience of complexity can arise just as powerfully in remarkably stripped-down contexts, extending even to very basic images that have few distinguishing features or pre-existing associations. For example, even the objects in Fig. 1B, which are just uniformly colored geometric shapes, give rise to a striking impression of complexity, with the shapes on the right again seeming more complex than the shapes on the left. What is the nature of this experience—and why do we have it in the first place?

Here, we explore these issues for the case of *visual objects*, a fundamental unit of perceptual processing. In particular, our experiments address three core questions about visual complexity (a "what" question, a "how" question, and a "why" question). First, we build on previous work asking what *about* visual objects make them look complex. Second, we ask how the mind extracts this property in the first place: Is complexity represented only through slower, deliberative judgments — or does it also arise in the course of fast, automatic visual processing? Third, we ask why our minds bother to represent complexity at all, and whether there are consequences of perceived complexity for other processes in the mind. We approach these questions by validating a new way to quantify the complexity of visual objects, and by examining the influence that complexity has on the deployment of voluntary and involuntary visual attention.

1.1. "What": Measuring object complexity

What *about* an object makes it look complex? The study of object complexity is not new (for a review, see Donderi, 2006), and indeed it may not seem particularly difficult to identify potential sources of complexity in an object or image. For example, in Fig. 1A, the complex images on the right contain more bumps, branches, and petals (respectively) than the images on the left. More generally, a natural thought is that visual

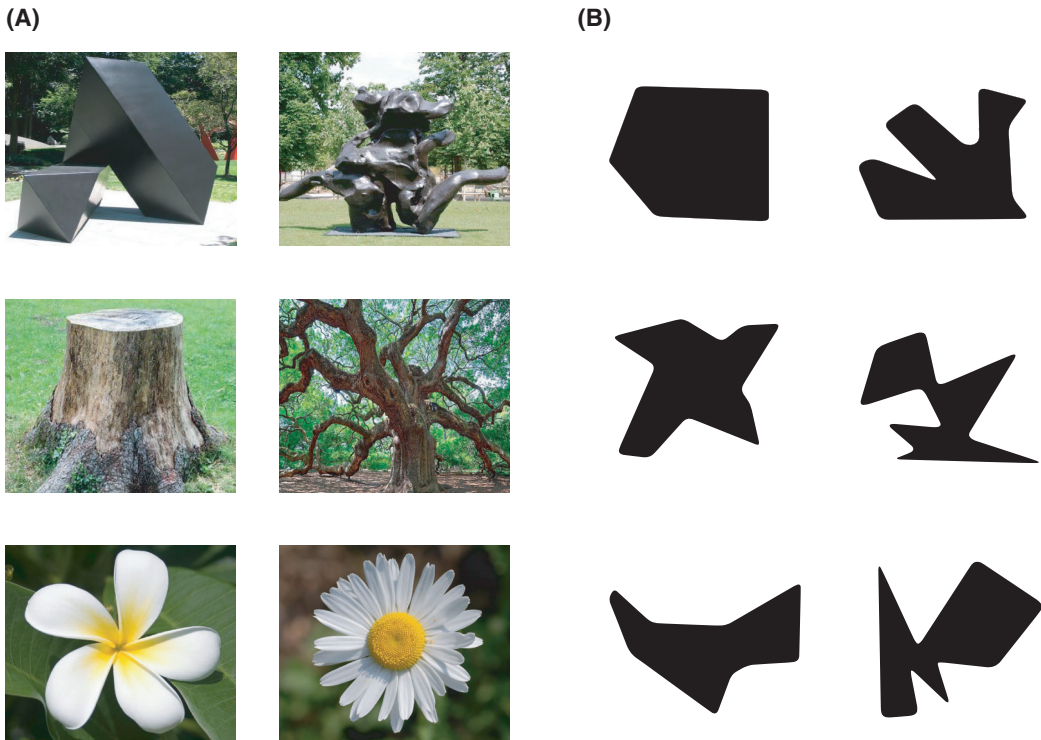


Fig. 1. Objects frequently strike us as being simple or complex. (A) The experience of complexity can span a surprisingly wide range of categories, shown here by examples drawn from both the natural and artificial world, as well as different materials, features, and scales. For each pair, the object on the right seems more complex than the one on the left. (B) The experience of complexity can arise even from basic shapes without many distinguishing features or pre-existing associations; here, too, the shapes on the right seem more complex than those on left. Moreover, this experience is not easily reducible to low-level features of the shape's bounding contours: The shapes in the second row, for example, have the same number of sides, or "turns"; and the shapes in the third row have the same number of sides and the same ratio of perimeter to area.

complexity is a function of the number of features present in a stimulus, where more features lead to greater complexity. Indeed, a seminal study of object complexity from the middle of last century showed that, for the sorts of objects in Fig. 1B, the key measure was the number of sides, or "turns," in the shape, which accounted for a majority of variance in explicit ratings of complexity (Attneave, 1957). Other features that have been identified include asymmetry, compactness (often computed as the ratio of squared perimeter to area), and even simply the amount of ink used to create the object, with more ink implying greater complexity (Attneave, 1957; Berlyne, 1958; Chipman & Mendelson, 1979; Cutting & Garvin, 1987; Day, 1968; Donderi, 2006; Heaps & Handel, 1999; Pelli, Burns, Farell, & Moore-Page, 2006; Watson, 2011).

However, one reason to suspect that *features* per se could not sufficiently account for perceived object complexity is simply that, for any given feature, it may be possible to

equate the presence of that feature across images and still observe differences in perceived complexity. For example, the second pair of shapes in Fig. 1B is equated for the number of turns in the shape; however, they still seem to markedly differ in their complexity. And the third pair of shapes in Fig. 1B is equated both for number of turns and perimeter-to-area ratio, but again they seem to easily produce a difference in subjective complexity.

In part for these reasons, an alternative approach to characterizing visual complexity is not to count up the features in an object, but instead to compute a summary statistic over the entire image. For example, more recent work has operationalized complexity in terms of the degree of visual clutter in an image (Rosenholtz, Li, & Nakano, 2007; see also Oliva, Mack, Shrestha, & Peeper, 2004), its self-similarity (Spehar, Walker, & Taylor, 2016), or even the size of the digital file in which the image is stored (e.g., JPEG file size; Forsythe, Nadal, Sheehy, Cela-Conde, & Sawey, 2011; Machado et al., 2015; Madan, Bayer, Gamer, Lonsdorf, & Sommer, 2017; Marin & Leder, 2013; Schlochtermeyer et al., 2013). The insight in this work is to connect perceived complexity to notions from information theory (Shannon, 1948), where complexity is inversely related to the “predictability” or “compressibility” of a given image (Attneave, 1954). For example, a larger file size often implies that there are fewer regularities across an image, which, in turn, implies greater complexity.

However, whereas these summary approaches are especially well tailored to visual *images* (e.g., natural scenes, or visual arrays composed of many elements), they are perhaps less appropriate for visual *objects*, which are segmented entities having insides, outsides, and parts in ways that scenes typically do not. For example, in Fig. 1B, it is not literally clutter that distinguishes the complex images on the right from the simple images on the left, all of which are simply individual bounded objects without any additional congesting features.

Can the strengths of these two approaches be combined? Recent work suggests that the progress made in quantifying the complexity of scenes (which captures computationally primitive patterns in such images) can also be achieved for visual objects themselves (in a way that respects their status as bounded *things* rather than mere collections of edges and features). In particular, objects can be described not only by their visible exteriors (Feldman & Singh, 2005) but also by their internal “skeletons”—inferred geometric structures based on axes of local symmetry (Blum, 1973; Feldman & Singh, 2006). The skeleton of a shape might be thought of as its “blueprint”: the internal structure that explains why a shape has the external features it does (Fig. 2A). Shape skeletons capture many aspects of a shape’s large- and small-scale organization, including not only the number of parts a shape has (Siddiqi, Tresness, & Kimia, 1996) but also how these parts are configured with respect to one another, and even to some extent the complexity of each part itself (since the shape of a skeletal branch captures the shape of its corresponding part). Shape skeletons are also psychologically plausible, with growing empirical evidence that they are computed and represented by human vision (Ayzenberg, Chen, Yousif, & Lourenco, 2019; Ayzenberg & Lourenco, 2019; Firestone & Scholl, 2014; Lowet, Firestone, & Scholl, 2018; Wilder, Feldman, & Singh, 2011).

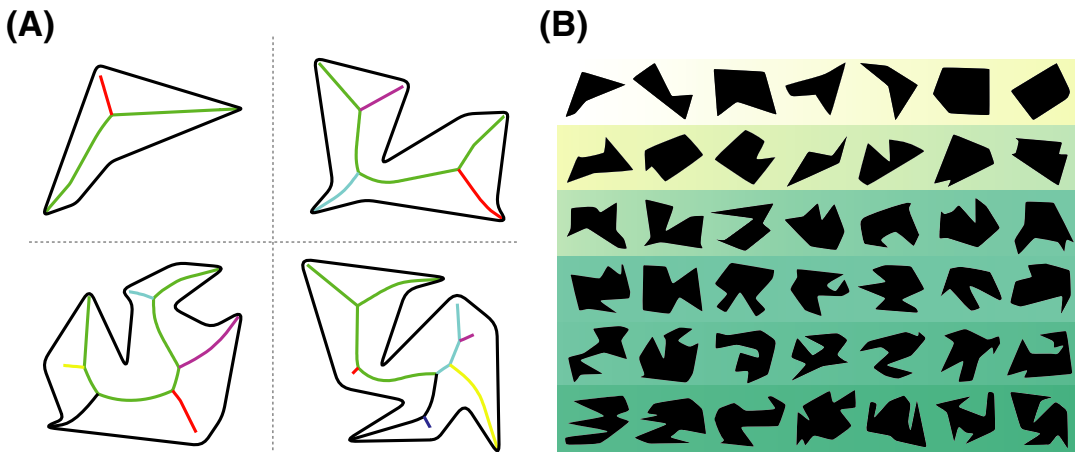


Fig. 2. Sample stimuli used in the present experiments. (A) Four shapes from the stimulus pool, with their medial-axis skeletons inset (in the actual experiments, these shapes appeared with black fills, and without their skeletons; the different colors are used to make the medial-axis branches more distinguishable, but are not significant in any other way). The measure of object complexity that we explore here is based on the complexity of this skeletal structure. (B) Sample objects from the library of 100 stimuli, ranked by structural surprisal (with simpler objects appearing at the top, and more complex objects appearing at the bottom).

A unifying approach, then, is to compute a measure of complexity over this internal structure, by asking how predictable or surprising a shape's skeletal description is. This approach, which we refer to here as “structural surprisal,” owes in recent years to Feldman and Singh (2006) and computes the integral of the turning angle along each skeletal branch; the complexity of a shape is then the cumulative surprisal of all the branches of its internal skeleton and reflects both how many branches a shape's skeleton has as well as how much each branch deviates from “smoothness.” Importantly, this approach takes seriously the objecthood of shapes (rather than treating them as any other form of visual image) while also allowing for the sort of rigorous computational analysis that has made information-theoretic approaches so attractive.

Indeed, recent work that explored questions other than perceived complexity per se found that skeletal structure of this sort predicts visual performance in other context (e.g., the ability to detect the outline of a shape when masked with noise; Wilder, Feldman, & Singh, 2016). We think the promise of this approach has yet to be fully realized, and in particular that it may help make progress on unanswered questions about how object complexity is extracted and represented in the mind.

1.2. “How”: Does complexity reflect seeing or thinking?

Over and above the question of how complexity is characterized, a perhaps even deeper question remains as to what kind of process in the mind is responsible for extracting the complexity of objects. In particular, determining the complexity of an image or object

might be the sort of process that is relatively slow, deliberate, and effortful, akin to looking at a shoe and deciding how expensive or popular it might be; and indeed, there is little doubt that judging complexity *can* proceed that way. But alternatively—or additionally—complexity might be also processed in a manner that is fast and automatic, akin to perceiving the shoe’s color or location.

Whereas nearly all of the scientific attention paid to visual complexity (including the work reviewed above) has focused on the *determinants* of visual complexity, these many and varied approaches are essentially silent on its *nature* in the mind. Do impressions of complexity reflect only thoughtful higher-level judgments, or are they also computed in the course of automatic visual processing? A central purpose of ours is to investigate this question empirically, by asking whether complexity is somehow privileged by core mechanisms of perception and attention.

1.3. “Why”: *Complexity and curiosity*

A final open question about the mind’s processing of complexity is why this impression arises in the first place. For basic visual properties such as color, size, or orientation, it is little mystery why the mind goes through the trouble of computing them, given their clear and critical survival benefits. For example, in detecting the ripeness of fruit or the approach of a dangerous predator, it is crucial to compute properties such as color and motion. But why bother representing the complexity of an object?

One possibility is that complexity promotes a kind of adaptive exploratory behavior, by arousing curiosity about relevant objects in our environment (Berlyne, 1960, 1966). Indeed, research in developmental psychology exploring the factors that encourage infants and children to interact with various objects has found that complex, irregular, and heterogeneous objects attract play and other exploratory behaviors, compared to simple, regular, and homogeneous objects—even when both the simple and complex objects are equally novel and unfamiliar (Switzky, Haywood, & Isett, 1974). Broadly similar findings have also been reported in adult observers, who spend more time looking at complex objects than simple ones (Berlyne, 1957, 1958, 1970). However, much of this classic research was conducted without any robust way to quantify complexity, relying instead on subjective impressions of Gestalt properties such as “incongruity” or “heterogeneity of elements” (and so instead often used the sorts of problematic features discussed above, such as the amount of ink taken to draw the object). More recent work in infants and young children has quantified complexity more rigorously by exploring the complexity of sequences of presented objects, as defined by the entropy or surprisal of their transition probabilities, and investigating their connection to looking time (Kidd, Piantadosi, & Aslin, 2012, 2014). But what of the complexity of single objects appearing alone? Could the surprisal of an object itself have similar consequences for exploration?

1.4. *The present studies*

The present research investigates the processing and function of object complexity. We algorithmically generated a library of shapes of the sort appearing in Fig. 1B and

determined their complexity by computing the surprisal of their internal skeletal structures. We then investigated (a) whether visual complexity is privileged in visual perception, as revealed by asymmetric visual search for complex objects versus simple objects (Experiments 1–3); (b) whether complex objects as defined by this approach promote a kind of exploratory engagement relative to simple objects (Experiments 4–5); and (c) whether this surprisal-based quantification predicts explicit and subjective judgments of complexity (Experiment 6). Collectively, our studies explore the utility of information-theoretic approaches for studying the nature of visual complexity.

2. Object generation and complexity quantification (for all experiments)

To establish a stimulus pool for all of the experiments reported below, we adopted an algorithmic procedure to generate a variety of “random”-looking shapes that obey certain constraints. (Given our aim to isolate *visual* complexity per se, we avoided familiar objects and shapes that have pre-existing associations and meanings; Snodgrass & Yuditsky, 1996; Op de Beeck & Wagemans, 2001; Wagemans et al., 2008). For each shape, we first defined the number of sides that the shape would ultimately have and then generated a set of randomly located points that would serve as the vertices of the shape’s edges. We then connected these points using the method of Delaunay Triangulation, which maximizes the minimum angles formed by the overall structure of the connected lines. Next, facets along the boundary of this triangle mesh were removed until the resulting polygon had the predefined number of sides. Finally, for each shape, the edges of the resulting polygon were smoothed in order to appear more natural, and the shapes were rendered as solid black objects on a white background. Additional constraints included a minimum angle no smaller than 20° , a maximum angle no greater than 160° , a minimum side length of 0.01 “units” (where 1 unit corresponds to the height and width of the plane in which the initial points were randomly distributed), and a minimum distance from any vertex to its nearest side greater than 0.05 “units” (to prevent shapes with unnaturally skinny “necks”; Siddiqi et al., 1996).

To calculate the “objective” complexity of each shape, we first derived its medial axis skeleton (using ShapeToolBox1.0; Feldman & Singh, 2006), classically defined as the set of points having two or more closest points on the shape’s perimeter (Blum, 1973). This operation returns a geometric structure comprising a series of connected branches, which form the basis for the measure of complexity that we use here (Fig. 2A). The present measure derives from earlier work in computational geometry (Feldman & Singh, 2006; see Wilder et al., 2016 for a recent application¹), and it reflects the integral of the turning angle along each skeletal branch, over the total number of branches. An intuitive way to capture this measure—which we refer to here as “structural surprisal”—might be to imagine a person walking along the skeleton of a shape; the more often this person changes direction (such that their next step was not easily predictable from their previous step), the higher the surprisal of their walk. The surprisal of a branch is thus its deviation from smoothness—more surprising branches are those with a greater tendency to zig and zag.

Cumulative surprisal is the sum of these zigs and zags, plus a penalty for the number of branches itself (so that shapes with more branches are more surprising, over and above the smoothness of each branch). What results is a measure of the “amount of information” required to describe the shape in terms of its basic underlying structure. The code for computing this measure is available at <https://osf.io/aev3j/> (along with all data, materials, code for stimulus generation, experiment code and [where applicable] pre-registrations of the present experiments).

3. Experiment 1: A search asymmetry for complexity

Even though a considerable amount of attention has been paid to the *determinants* of object complexity, nearly all previous work on perceived complexity (including all of the work reviewed earlier) is silent on the question of *what kind* of process in the mind extracts the complexity of an object. In particular, is object complexity the sort of property that requires deliberation and reflection to arise, as when we consider how expensive or fashionable an object is? Or can it be more like those properties that are *seen* directly and with minimal effort, such as an object’s size or orientation? To our knowledge, this question has not previously been asked about the perception of complexity.

To investigate this question, we asked whether complexity can serve as a target for visual search. Visual search refers both to our ability to find specific targets within a broader array of distractors, and also to a popular psychophysical approach to investigate which features may be prioritized or privileged by visual processing (Treisman & Gelade, 1980; Treisman & Gormican, 1988; Wolfe, 2001). In particular, features which are prioritized by visual processing typically exhibit a phenomenon known as *search asymmetry*, whereby the time taken to find an object defined by a given feature (among distractors that do not have the feature) is less than the time taken to find an object without that feature (among distractors that do have the feature). For example, it may be easier to find a cross (“+”) among vertical lines (“|”) than it is to find a vertical line among crosses, because crosses have a feature—an intersection—that vertical lines do not. Search asymmetries can also arise for properties whose spaces are organized more continuously: For example, size is a factor that “undoubtedly” guides visual search (Wolfe & Horowitz, 2017); however, unlike intersections, *being big* is not naturally described as a property that objects either have or do not have. Instead, every object is some size; and differences in size, if large enough, can facilitate search behavior.

Here, we ask whether complexity is prioritized in this way. Using the sorts of objects created above (Fig. 2), we paired objects from the “complex” end of the library with objects from the “simple” end of the library and manipulated whether the complex objects were targets among simple distractors or the simple objects were targets among complex distractors. If complexity can be processed more like size and orientation than like expensiveness or popularity, then complex objects should be easier to find among simple distractors than vice versa.

3.1. Method

3.1.1. Subjects

12 students from Johns Hopkins University completed the experiment in exchange for course credit.

3.1.2. Apparatus

Stimuli were presented on a BenQ Zowie XL LCD Display (24.5 inch) with a 60Hz refresh rate, using custom software written in Python with the PsychoPy libraries (Peirce, 2007). Subjects sat approximately 57 cm from the display.

3.1.3. Stimuli

100 objects were generated using the procedure mentioned in the previous section, ranging from 5-sided objects to 14-sided objects, with 10 instances of each N-sided object (Fig. 2B). 40 complex objects and 40 simple objects were selected from the extrema of the complexity spectrum, according to their structural surprisal. Each shape was approximately 4 cm × 3 cm as presented on the display.

3.1.4. Procedure

In each trial of the experiment, the subject's task was simply to indicate the presence or absence of an "oddball" item within a larger display; in other words, to indicate whether all the items were the same, or whether one was different from the rest.

Each trial of the experiment fell into one of four conditions (Fig. 3): (a) *Complex Target*, in which a single complex item appeared among simple distractors that were all identical to each other; (b) *Simple Target*, in which a single simple item appeared among identical complex distractors; (c) *All Complex*, in which all the items were the same complex object, with no simple target present; and (d) *All Simple*, in which all the items were the same simple object, with no complex target present.

The set size of the display (6, 12, or 18 items) was also crossed with these four conditions, yielding 4 (conditions) × 3 (set sizes) = 12 unique trial types. All 40 simple-complex object pairs appeared as each of these trial types, for 12 × 40 = 480 trials total.

Each trial began with a fixation cross at the center of the display for 500 ms. Immediately after the fixation cross disappeared, the objects for that trial were shown on a gray background, at random positions within an invisible 6 × 6 grid. To best compare the search performance across trial types, every group of trials involving the same simple-complex object pairs at the same set size used the same locations within the grid (with only the identities of the target and distractor objects exchanged).

To indicate the presence of a target, subjects pressed the "S" key; to indicate the absence of a target, subjects pressed the "L" key. Subjects were also instructed to respond as quickly as possible without sacrificing accuracy (though there were no enforced constraints on response times). After a response, the fixation cross for the next trial appeared.

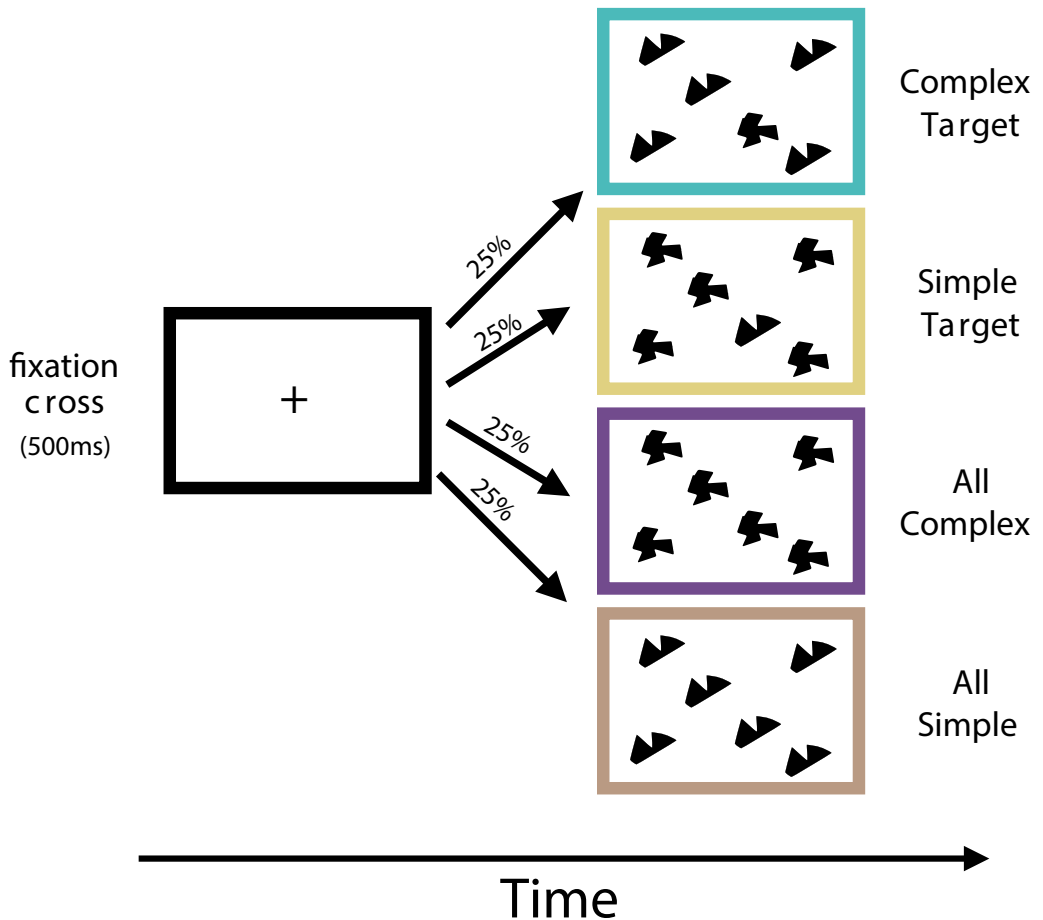


Fig. 3. Illustration of the four primary conditions used in Experiment 1's visual search task (showing set size 6 as an example). On each trial, a fixation cross appeared for 500 ms, followed by a display from one of the four conditions. The subject's task was to answer whether all the objects were the same, or whether one object differed from the rest.

3.2. Results and discussion

As expected, subjects had little difficulty completing the task, with a mean accuracy of 95% and a mean response time (across all conditions) of 900 ms.

Crucially, subjects were faster to find a complex target among simple distractors than vice versa (790 ms vs. 857 ms; $t(11) = 8.00$, $p < .001$, $d = 2.31$; effect sizes here are computed as the mean RT difference between conditions, divided by the standard deviation of the within-subject differences between condition means; Fig. 4). These results revealed a classic search asymmetry, here caused by the presence and absence of complex targets. Moreover, subjects were also more accurate at finding complex targets than simple targets (95% vs. 90%; $t(11) = 2.29$, $p < .05$, $d = .66$); this provided further

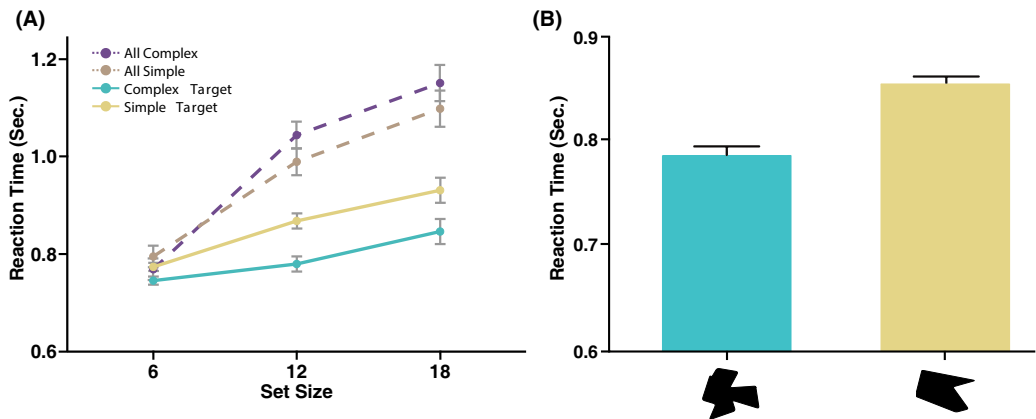


Fig. 4. Performance on the visual search task in Experiment 1. (A) Finding a complex target among simple distractors proceeds efficiently (teal line; slope = 8 ms/item). (B) Comparing the average search time between the “complex target” condition and the “simple target” condition shows that finding a complex object among simple objects is faster than vice versa ($p < .001$), a classic search asymmetry. Error bars depict ± 1 SEM of the difference between conditions.

evidence of asymmetric processing of complexity, and it also suggests that the search asymmetry in reaction time did not reflect a speed-accuracy trade-off (since subjects were both faster *and* more accurate at detecting complex targets).

Moreover, we also observed evidence that complex targets were found “efficiently”, even beyond the asymmetry in search performance. Classically, a search asymmetry is considered evidence of efficient processing when search slopes are relatively flat; for example, search slopes for oddball visual features are typically < 10 ms/item (Wolfe, 1998). We observed this pattern of efficient search for complex objects: Considering the different set sizes at which simple distractors appeared, the reaction-time cost for each additional distractor was only 8 ms/item (calculated here, as in other work, by taking the difference in reaction time between set size 6 and set size 18 and dividing that difference by the difference in set size itself [here, 12]); by contrast, search for simple targets was not “efficient” according to these traditional criteria, with a slope of 13 ms/item (with this slope difference itself being marginally different from the slope for complex targets; $F(2, 22) = 2.96$, $p = 0.07$, for the interaction between set size [6, 12, or 18] and target type [complex or simple]; Fig. 4).

On one hand, we do not read too deeply into this last result, in part because there is an active debate in the literature over exactly what constitutes efficient processing and what this implies about the mind more generally (Rauschenberger & Yantis, 2006; Rosenholtz, 2001; Wolfe & Utochkin, 2019). But on the other hand, we took the entirety of these results as encouraging initial evidence for our broader hypothesis that representations of complexity arise not only through slow, deliberative judgmental processes but also through processes that proceed more rapidly and with relatively minimal effort.

4. Experiment 2: Complexity per se

The previous experiment found that complex objects were easier to visually locate than simple objects, suggesting that complexity may be subject to the relatively basic processing constraints of visual cognition. However, complexity as measured here by structural surprisal tends to correlate with other low-level properties of shapes; in that case, the search asymmetry observed in Experiment 1 may have been driven merely by such low-level properties, rather than by more sophisticated representations of complexity itself. For example, the complex shapes in Experiment 1 tended to have more sides than the simple shapes, and it is possible that shapes with more sides are easier to find regardless of how complex they are in any deeper sense.

Even though such low-level visual factors may well be *cues* to complexity (e.g., as part of an “extraction algorithm”; Halberda, 2019), Experiment 2 further isolated our computationally more basic notion of complexity by asking whether complex targets drive faster visual search even when important low-level shape properties are equated across simple and complex objects. We generated a new library of shapes that were matched for number of sides, the most prominent contributor to judgments of visual complexity in previous work (Attneave, 1957; Day, 1968). By doing so, we also controlled for several other low-level properties, including area, spatial frequency, various angular properties, and others (see Section 4.1 for more details).

In this way, we explore whether the mind’s computation of visual complexity goes beyond certain salient low-level features and instead represents a deeper property of objects, by asking whether a search asymmetry exists even when such features are equated.

4.1. Method

4.1.1. Subjects

12 new students from Johns Hopkins University completed the experiment in exchange for course credit.

4.1.2. Stimuli and procedure

A new library of objects was generated using the same procedure as in Experiment 1, except that this time the shapes were chosen in advance to have either 10 or 12 sides (Fig. 5A). We selected 25 complex objects and 25 simple objects from the complexity extrema of a spectrum of 10-sided shapes, and another 25 complex objects and 25 simple objects from the complexity extrema of a spectrum of 12-sided shapes, for a total of 100 simple and complex shapes. Each of the 25 complex shapes was initially paired with its corresponding simple shape at the other end of its spectrum, to create 50 shape pairs (i.e., 10-sided complex shapes paired with 10-sided simple shapes, and 12-sided complex shapes paired with 12-sided simple shapes); then, we also paired the 10-sided shapes with the 12-sided shapes, to create another 50 pairs (i.e., 12-sided complex shapes paired with

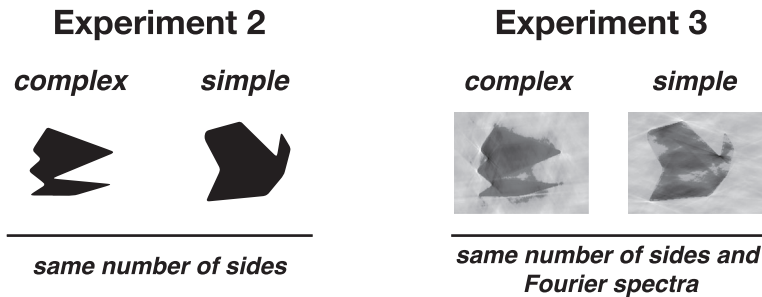


Fig. 5. In Experiment 2, complex objects and simple objects were equated for number of sides, or “turns.” In Experiment 3, the same complex objects and simple objects from Experiment 2 were further equated using a phase-matching procedure.

10-sided simple shapes, and 10-sided complex shapes paired with 12-sided simple shapes), for a total of 100 shape pairs.

Incidentally, controlling for number of sides also ended up equating many other properties as well. Overall, the complex shapes and the simple shapes differed by only 0.8% in terms of convex hull area, 0.4% in terms of spatial frequency, and 1.9% in terms of the range of angles present in the shape. Though these properties were not literally identical, these very small differences encouraged us that any larger differences in search speed were unlikely to be driven by, for example, differences in convex hull area or spatial frequency.

Otherwise, the conditions in the experiment were the same as in Experiment 1: (a) Complex Target; (b) Simple Target; (c) All Complex; (d) All Simple. All 100 simple-complex object pairs appeared once in each of these conditions, always within a set of 18 items, resulting in a total of $4 \times 100 = 400$ trials. The experimental procedure itself was identical to Experiment 1.

4.2. Results and discussion

Once again, complex targets were found faster than simple targets (1,156 ms vs. 1,275 ms, $t(11) = 4.50$, $p < .001$, $d = 1.30$; Fig. 6),² even when simple and complex objects were precisely equated (or highly similar) on a range of low-level features. Additionally, subjects exhibited greater accuracy in finding complex targets than simple ones (93% vs. 88%; $t(11) = 3.24$, $p < .01$, $d = 0.94$), which again suggested that the reaction-time advantage was not the result of a speed-accuracy trade-off.

These results suggest that the visual system’s representation of complexity goes beyond various low-level features that may be correlated with complexity. Indeed, Attneave (1957) found that the number of sides (or number of “turns”) could account for most of the variance in subjective ratings of complexity. However, the present results suggest that visual processing also represents and prioritizes a deeper notion of complexity—here reflected in facilitated search for complex objects whose complexity goes beyond their mere number of sides (and other features).

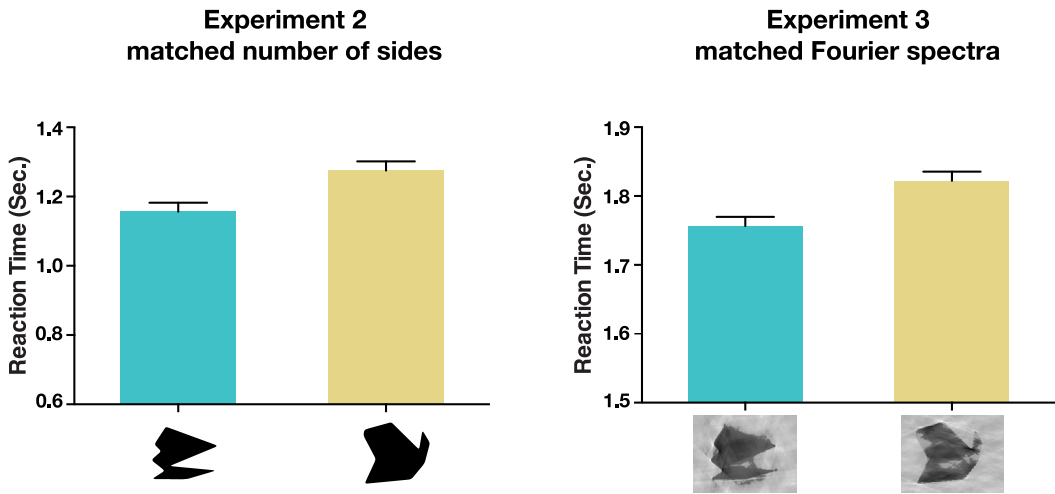


Fig. 6. Search times for complex targets versus simple targets in Experiments 2 and 3. In Experiment 2 (left), object pairs had the same average number of sides, or “turns.” In Experiment 3, the same objects were matched not only for those properties but also for their Fourier spectra. In both Experiment 2 and Experiment 3, complex targets were found faster than similar targets, replicating the search asymmetry from Experiment 1. Error bars depict ± 1 SEM.

5. Experiment 3: Complexity per se—Really!

Though the complex and simple objects in Experiment 2 were equated for number of sides and other features, a further worry may arise that *some* important low-level difference remains (even if it is difficult to precisely express what that difference might be), and that this difference drives faster search times for complex objects. VanRullen (2006) reported that equating the global statistical properties of images eliminated “pop-out” effects for faces over non-faces (though in a different design than in the present experiments; Hershler & Hochstein, 2005) and suggested “phase matching” as a way to isolate higher-level properties of interest. Although we do not explicitly argue here for complexity pop-out—only an asymmetry of search—we nevertheless implemented this suggestion to further explore the possibility that complexity per se is privileged by visual attention.

Experiment 3 used the same design and stimuli as in Experiment 2, but with the additional step of first passing all object pairs through a phase-matching procedure that not only equated the properties discussed in our earlier experiments but also equated the Fourier spectra for each pair of objects (which, as a consequence, equated the shapes on properties such as spatial frequency and luminance). This experiment thus applies what is perhaps the most stringent test recognized in the visual cognition literature for equating low-level image properties, and thereby isolates the higher-level property of visual complexity. Does the search advantage for complex objects persist even under these conditions?

5.1. Method

25 students from Johns Hopkins University completed the experiment in exchange for course credit. The same stimulus set in Experiment 2 was used in this visual search task, except that every object pair was matched using the SHINE toolbox (Willenbockel et al., 2010), which equates the Fourier spectra of two or more objects. For a large set of objects, this procedure identifies the average frequency spectrum of the set and uses an inverse fast Fourier transform to convert these matched frequencies back into the spatial domain. Additionally, this procedure also matched the luminance histograms of the set of objects, with the average histogram serving as the target luminance. Here, we applied this matching procedure independently for each pair of shapes, such that each complex-target/simple-distractor trial and its corresponding simple-target/complex-distractor trial were matched on these properties. For examples of the images it outputs, see Fig. 5B; as before, all of the images used here, as well as the code for producing them, are available at <https://osf.io/aev3j/>.

The design was the same as in Experiment 3, except that each subject completed half as many trials as in that experiment, with each subject's 200 trials randomly chosen from the broader pool of 400 possible trials.³ This design, including the sample size and analysis plan, was pre-registered before any data were collected (<https://aspredicted.org/9up4u.pdf>). (One subject was excluded for failing to reach the pre-registered performance cutoff; however, no pattern reported below depended on excluding this subject.)

5.2. Results and discussion

Overall accuracy across all trials was 90%. Crucially, despite being matched on a comprehensive list of low-level visual cues, subjects were again faster to find a complex target among simple distractors than vice versa (1,757 ms vs. 1,822 ms, $t(23) = 2.28$, $p < .05$, $d = .47$; Fig. 6). Moreover, this advantage was not explained by a speed-accuracy tradeoff, since subjects were more accurate on complex-target trials than simple-target trials (83% vs. 80%; $t(23) = 2.81$, $p < .01$, $d = 0.57$).

This result suggests that subjects' search behavior is truly responsive to complexity per se, which can emerge over and above various low-level visual properties that may be correlated with complexity.

6. Experiment 4: Complexity promotes exploratory engagement

The preceding experiments have explored the possibility that object complexity is computed by the mind in ways that can guide the deployment of visual attention, over and above low-level properties that may be correlated with complexity. But these results invite another question: *Why* is complexity a salient feature in visual perception? In other words, why does the mind bother to compute complexity in the first place?

For other core visual properties such as color or shape, it is rather unmythical why the visual system might compute them. Color, for example, can be a signal of food

ripeness, facial emotions, toxic substances, etc., and so our perception of color can function to inform us of survival-relevant factors in our environment. But what is the function of perceived complexity?

More than half a century ago, Berlyne (1966) noticed that people invest considerable time and energy seeking out stimuli that are not inherently beneficial for survival and suggested that stimuli that are more complex may trigger stronger “curiosity.” Indeed, empirical work on this question has revealed that adults (Berlyne, 1957; Lewis & Frank, 2016) and children (Switzky et al., 1974) look at and explore complex objects and sequences more than simple ones. This insight suggests that one function of complexity might be to promote exploratory behavior toward such stimuli, perhaps because complexity may be an external signal that an object has more to reveal, or perhaps to reduce the uncertainty associated with irregular or unpredictable objects and events (e.g., to find a simpler description of a superficially complex object). However, much of this earlier work, though suggestive of this possibility, used highly heterogeneous stimuli whose complexity cannot be varied in an objectively quantifiable way, and even introduced new confounding factors, leaving the precise role of object complexity in exploration unclear.

Here, we investigated the connection between complexity and exploration using the objective measures explored in our above studies. We ask whether complexity, as quantified by structural surprisal, promotes a kind of attentional engagement with objects in an unconstrained task that asks subjects to interact with stimuli as they please.

6.1. Method

6.1.1. Subjects

120 subjects completed this experiment through Amazon Mechanical Turk and were monetarily reimbursed. (For a discussion about the reliability of this subject pool, see Crump, McDonnell, and Gureckis [2013].)

6.1.2. Stimuli

Using the procedure outlined above, we generated a much larger pool of 2,000 objects and selected the 50 most complex shapes and 50 most simple shapes from the extrema of this pool. (Normalized structural surprisal: $M = 0.89$, $SD = 0.04$ for complex shapes; and $M = 0.11$, $SD = 0.03$ for simple shapes.)

6.1.3. Procedure

Subjects were shown all 100 shapes, one at a time in a random order, and were told to study each of them for a later memory test. After every 20th object, subjects were shown four “test” objects and were asked whether they had seen these objects before; in fact, two of the shapes were always drawn from the block of shapes that the subjects had just seen, and two were drawn from a similar pool of shapes that never appeared anywhere else in the experiment.

The key feature of this design, inspired by a recent study on word length and object complexity (Lewis & Frank, 2016), was its self-paced nature: Subjects were permitted to study each shape as long as they liked, and they pressed a key when they felt ready to move on to the next object. The experience for subjects was thus akin to “browsing” the entire library of shapes, freely advancing or pausing as they wished. Our measure of interest was not their actual memory performance, but rather how long they chose to view each object. Readers can experience this task for themselves at <https://perceptionresearch.org/complexity/exploration>.

At the subject level, we excluded anyone who failed to perform above 50% accuracy in the recall session. At the trial level, we excluded trials with viewing durations greater than 10 s or shorter than 500 ms (1.9% of all trials). These criteria excluded 16 subjects, leaving 104 subjects with analyzable data. (No pattern reported below depended on excluding these subjects; i.e., all effects remained significant, in the same direction, even when these subjects were included in the analysis.)

6.2. Results and discussion

The recall task was moderately difficult for subjects, with an average accuracy for recalling objects of 78% (82% for simple, and 74% for complex). Our primary interest, however, was in the self-chosen study times for each object. We observed that subjects spent a longer time studying complex objects than studying simple objects (2.09 s vs. 1.71 s, $t(103) = 7.25$, $p < .001$, $d = 0.70$; Fig. 7).

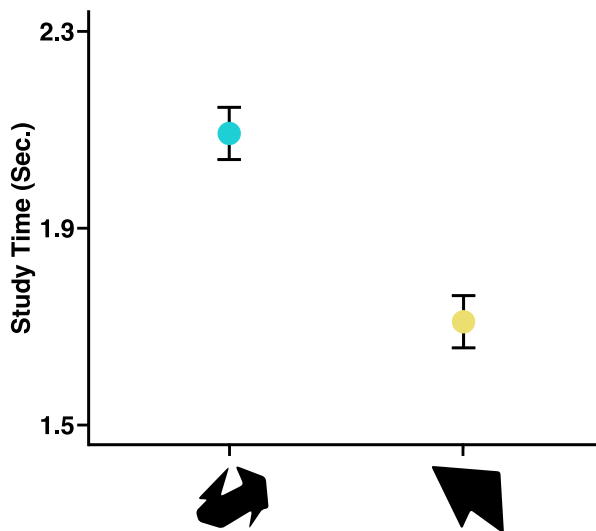


Fig. 7. Performance on the self-paced study task of Experiment 4. Subjects spent a longer time studying complex objects when they were asked to later recall them. Error bars depict ± 1 SEM.

This result thus provided initial evidence that complexity promotes a kind of exploratory attentional engagement: Subjects spent a longer time engaged with complex objects than they did with simple objects.

7. Experiment 5: Enhanced exploration even when complexity is incidental

The previous experiment showed that subjects study complex objects for a longer duration of time than simple objects, suggesting a connection between complexity and exploration or attentional engagement. However, this may well have been an explicit strategy on the part of subjects: Subjects may have consciously chosen to spend more time studying complex objects when being tested on those objects later on—especially since complexity of the shapes was, at least indirectly, the very same property that was being tested.⁴ Indeed, this was also true in a previous investigation of study time for objects, which found that the number of geons in an object predicted increased study times in a similar task (Lewis & Frank, 2016). However, if complexity is truly connected to engagement, then might it be powerful enough to exert its influence even when shape complexity is completely irrelevant to the task? To our knowledge, no prior study has asked this question.

Whereas Experiment 4 may have “stacked the deck” in favor of our hypothesis, Experiment 5 stacked the deck *against* our hypothesis by making the complexity of the objects, and indeed even their shapes as a whole, completely incidental to the task. As in Experiment 4, subjects in Experiment 5 saw sequences of objects and were later tested on them. However, they were not asked to recall the *shapes* of the objects; instead, they were asked to recall the *texture* that the objects happened to be rendered in. If complexity is computed by the mind in ways we cannot fully ignore or disregard, then perhaps we could observe a (subtle) behavioral influence even when it is completely task-irrelevant.

7.1. Method

7.1.1. Subjects

A pilot study suggested that a sample size of 300 subjects would achieve 95% power to find an effect of complexity on exploration of textured shapes. We thus pre-registered this sample size, as well as an analysis plan, which was the same as in Experiment 4 (<https://aspredicted.org/y6nm8.pdf>).

7.1.2. Stimuli

In addition to using simple and complex shapes, we also designed 20 textures (e.g., stripes, diamonds, zig-zags, and other patterns) that were each rendered in five different colors (blue, green, pink, orange, purple), making 100 unique textures that were used as “fills” for the objects. The 100 objects used in this experiment were the same as Experiment 4, except for the addition of these fills. Each shape was presented in a unique fill

throughout the experiment, with fills and shapes matched up in a different random pairing for each subject.

7.1.3. Procedure

As in Experiment 4, subjects saw objects presented one after the other, with the next object appearing when the subject pressed a key. This time, however, subjects were tested only on the fills inside each object, rather than on the shapes of the objects (and were told this in advance). In the recall session, subjects were shown four textured squares and had to indicate which square had a texture that they had seen before (with two drawn from the block that they had just viewed and two being novel). The subject's task was to indicate, for each texture, whether it had appeared in the previous block. Thus, the actual shapes of the objects could be completely ignored, since they were irrelevant to performance on the recall task.

7.2. Results and discussion

As expected, this task was especially difficult (since subjects had to recall not only the pattern on the objects but also that pattern's color), and so our pre-registered analysis plan conservatively excluded approximately one-third of subjects: 31% of subjects were excluded for failing to perform above chance on the recognition task. For those 300 subjects included in our analysis, overall accuracy in recalling textures was 71% (70% for simple, and 71% for complex), and 2.4% of trials were excluded because the viewing duration was greater than 10s. (No trial was excluded for being too fast.) Remarkably, however, even though the nature of the recall task made the shapes of the objects (and therefore their complexity) completely task-irrelevant, subjects again looked at complex objects for a longer time than simple ones (1.81 s vs. 1.78 s, $t(299) = 2.70$, $p < .01$, $d = 0.16$; Fig. 8).

This difference, though small in its raw magnitude, implies that complexity may be a powerful enough promoter of exploration that it exerts its influence even when subjects can and should be actively ignoring it. Even when complexity was completely irrelevant to the subjects' task, they continued to view complex objects for a longer duration than simple objects.

8. Experiment 6: Judging complexity

The previous experiments used various implicit measures to explore how the mind processes the complexity of visual objects, showing how object complexity guides visual attention and exploration. However, all of this work assumes that what we have referred to as "complex" and "simple" objects are actually perceived that way by subjects; in other words, whereas we used an "objective" measure to create and organize our stimuli, we have not yet confirmed that "subjective" measures would rank the objects similarly. Though we assumed this in part based on our own intuitions and impressions (e.g., upon

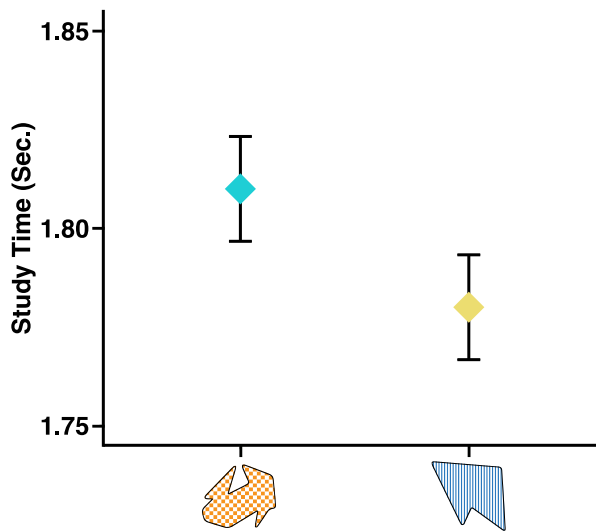


Fig. 8. Performance on the self-paced study task of Experiment 5. Subjects spent a longer time studying complex objects when they were asked to recall the shape's textures. Error bars depict ± 1 SEM.

looking at the array of shapes in Fig. 2B, which are arranged by their structural surprisal), our final experiment verifies these intuitions experimentally, by asking whether objectively more “surprising” objects also *appear* more complex to subjects.

8.1. Method

8.1.1. Subjects

50 subjects participated in this experiment through Amazon Mechanical Turk and were monetarily compensated for their participation.

8.1.2. Stimuli

All 100 objects generated for Experiment 1 were used in this experiment.

8.1.3. Procedure

All 100 stimuli were displayed to each subject, serially, in a different random order for each subject. For each object, subjects were asked to rate how “simple” or “complex” the shape looked, without any further explanation of how to make such a determination. Beneath each shape, a scale from 0 (simple) to 100 (complex) appeared, and the subject could move a slider along the scale to give their rating, with no time pressure. Once the subject was satisfied with the rating, they clicked a button to proceed to the next trial, at which point a different shape appeared on the display.

Readers can experience this task for themselves at <https://perceptionresearch.org/complexity/ratings>.

8.2. Results and discussion

To examine the relationship between structural surprisal and subjective impressions of complexity, we correlated each shape's computed surprisal value and the mean complexity rating across all subjects. Ratings of complexity were well distributed throughout the range of the scale, with a mean rating of 50.25 across all shapes and all subjects, and a standard deviation of 18.31. Crucially, structural surprisal was highly correlated with subjective complexity: Even at the level of the individual shapes themselves, objects with a greater structural surprisal value were rated as more complex ($r(98) = .75, p < .001$; Fig. 9). Thus, the shapes we have called "complex" and "simple" in our previous experiments are indeed perceived that way by subjects.

Of course, even though structural surprisal predicted explicit complexity ratings here in Experiment 6, it is nevertheless possible that complexity ratings (including those we

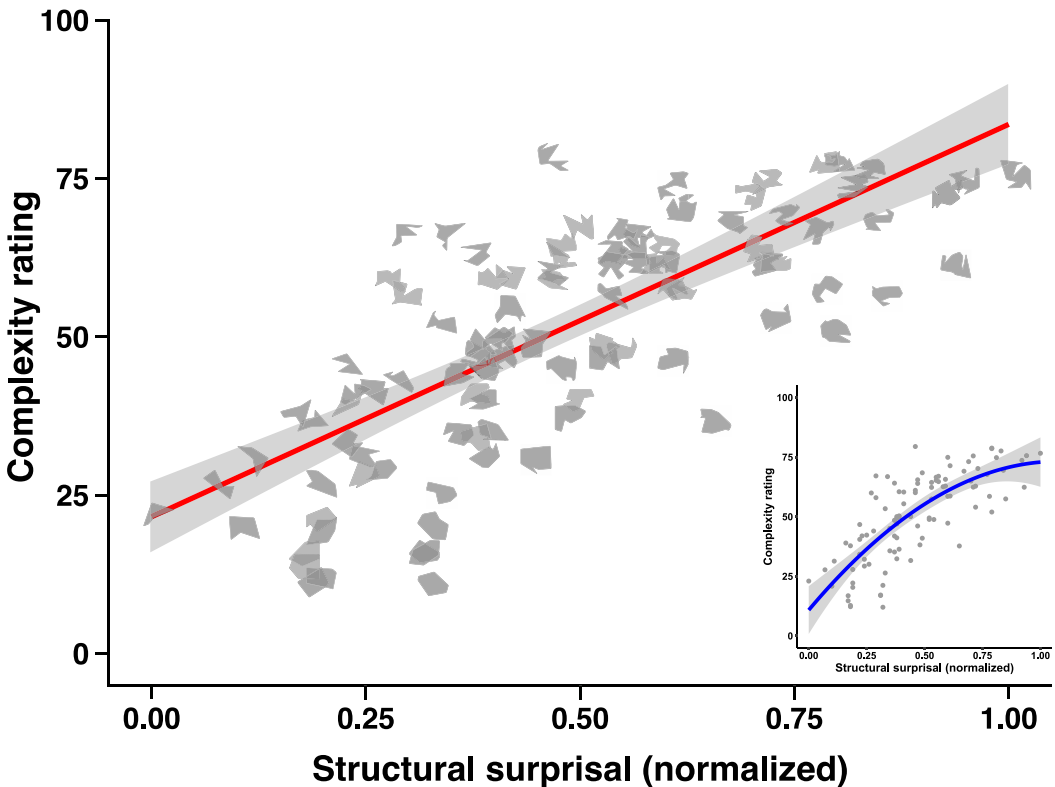


Fig. 9. The relationship between explicit complexity ratings for a shape (y-axis) and that shape's structural surprisal (x-axis). Each point is represented using the actual shape shown in the experiment. As structural surprisal increases, so do judgments of perceived complexity. However, we also note that these data might be modeled non-linearly (e.g., the quadratic fit shown in the bottom right panel); future work may further probe the relationship between the "objective" complexity of an image or object and the "subjective" impressions of complexity that subjects form upon seeing it.

collect here) have a different cognitive origin than the effects observed in our previous experiments. For example, subjective judgments of complexity could rely on an explicit, high-level, reflective processes, while attentional search and exploration rely on more implicit, automatic visual processes. Though parsimony perhaps favors a single process to explain these otherwise coincidental patterns, we acknowledge that the self-report measure itself will not be able to rule out this possibility. An intriguing question for future work, then, is whether object complexity is one thing or many, or whether different “parts” of the mind compute different representations of complexity.

9. General discussion

The six experiments reported here explore the representation and function of visual complexity. First, we discovered that complexity guides visual attention to complex objects such that they stand out among simple distractors in visual search arrays (Experiment 1). This search asymmetry persisted even when other properties were equated, including not only contour-based properties such as number of sides, area, and range of angles (Experiment 2) but also image-based properties such as spatial frequency and even the Fourier spectra of the targets and distractors (Experiment 3). Second, self-paced interaction with simple and complex objects revealed that visual complexity promotes engagement with stimuli in one’s environment: Subjects chose to view complex objects for a longer duration than simple objects (Experiment 4), even when the shape of the objects was completely irrelevant to the subjects’ task goals (Experiment 5)—implying that complexity stimulated a kind of curiosity about objects that encouraged subsequent exploratory behavior. Finally, we validated the surprisal of an object’s inferred interior structure as a measure of visual complexity by showing that it predicts not only implicit attention and exploration but also more deliberate judgments of complexity (Experiment 6).

Collectively, this work not only sheds light on the nature and function of visual complexity but also serves as a case study for how information-theoretic approaches can illuminate questions that had otherwise been surprisingly difficult to investigate in a computationally and psychophysically rigorous way.

9.1. A “blueprint” for complexity

Our particular approach here characterizes the complexity of a visual object in terms of the underlying entropy or “surprisal” inherent in it. However, there have been many other investigations of visual complexity that have focused on different parameters and properties (for a review, see Donderi, 2006); are these different approaches compatible with one another? Here, we do not intend to tease apart structural surprisal from other properties that may contribute to object complexity; indeed, there *must* be other factors contributing to perceived complexity that are simply not captured by the present measure (such as color and other non-geometric features). But we do think the present notion of surprisal or unpredictability combines the strengths of prior approaches while improving

on their potential weaknesses, and perhaps even subsumes such approaches into a more unified account of what makes an object look complex.

Some prior approaches have sought to capture the complexity of objects by reference to quantifiable and “objective” features of their contours—for example, the number of sides a shape has (Attneave, 1957), the variance in its sides and angles (Mavrides & Brown, 1969), or the ratio of perimeter to area (Pelli et al., 2006; Watson, 2011). These measures often reflect intuitive judgments of complexity (e.g., more features = more complex), but we note that it is often possible to equate these various contour-based features while still observing robust differences in complexity (as in our Fig. 1B, as well as Experiments 2 and 3). This suggests that contour-based features may not ultimately be the most ideal approach to characterizing visual complexity.

By contrast, other approaches have appealed to more intuitive notions of complexity, based on properties such as “heterogeneity of elements” and “irregularity of shape” (Berlyne, 1958, 1970). These sorts of approaches seem to better capture visual complexity as being bound up with the organization or structure of an object; however, this more intuitive approach has struggled to become computationally precise. How should we define notions of “irregularity” and “heterogeneity” such that we can apply them objectively and precisely to new objects and situations?

The notion of structural surprisal, it seems to us, bridges these two approaches in a way that is both qualitatively intuitive and quantitatively precise. Appealing to a shape’s skeletal structure easily accommodates, and perhaps even explains, previous findings concerning the influence of number of sides and sharpness of angles on complexity judgments—since those sorts of factors are very often disposed to increase skeletal complexity. At the same time, structural surprisal provides a rigorous computational framework in which to understand more subjective or intuitive notions such as “heterogeneity,” “irregularity,” “incongruity,” etc., by cashing out such notions in terms of an object’s unpredictability or deviation from smoothness.

Similar considerations may apply to previous studies of “parthood” in object representations. A great deal of rigorous and creative work explores how the mind carves up objects into parts (Singh & Hoffman, 2001), identifying factors such as the prevalence of concavities in an object as a driver of such part-based decomposition (De Winter & Wagemans, 2006) and its influence on visual attention (Hulleman, Te Winkel, & Boselie, 2000). However, these studies are less concerned with perceived complexity per se (beyond the number and location of an object’s parts), and indeed there are considerable advantages to characterizing complexity in terms of surprisal, which captures not only the number of parts in an object but also provides a framework in which to *rank* the complexity of objects in ways that part-based analyses perhaps cannot. For example, is an object with one big concavity more complex or less complex than an object with two smaller concavities? Merely identifying the concavities themselves fails to answer this question, because it fails to quantify the strength or influence of such properties on perceived complexity; by contrast, the surprisal-based notion explored here not only identifies such parts (roughly corresponding to branches of the skeleton) but also allows for quantifying and comparing their contributions to perceived complexity.⁵

Indeed, the approach taken here for the complexity of objects seems analogous to other approaches that have recently been taken for sequences of events. Recent work, for example, has characterized the complexity of a stream of images in terms of the transitional probability of the images (Kidd, Piantadosi, & Aslin, 2012), with the complexity of a sequence corresponding to its predictability. The present work shows how the complexity of visual objects can be understood in a similar way: Just as this other work asks how predictable the next image is given the images one has seen so far, structural surprisal might be thought of as asking how predictable some region of an object is from nearby regions. This suggests that predictability of this sort may be a generally applicable approach to characterizing complexity, beyond the particular cases we consider here.

9.2. *Broader notions of complexity*

Of course, the subjective experience of complexity, including visual complexity, can be found all around us: As shown in Fig. 1, a strikingly wide range of visual categories can be experienced as simple or complex, including—as explored here—basic geometric shapes of the sort that a child might draw on a piece of paper. The content of this fuller class of complex experiences is extremely rich, and it goes far beyond the geometric shapes used in the present work. It is worth commenting on this disparity: Do basic blob-like shapes really inform us about complexity writ large?

Though a complete theory of visual complexity must ultimately grapple with such broader notions and experiences (and this is an opportunity for future work), we think that stimuli like those used here are in fact perfectly suited to this initial investigation of visual complexity, for at least three reasons. First, a major advantage of these stimuli is the rigorous framework in which they can be generated and characterized. In the present work, this allowed us to study their complexity as distinct from so many other confounding factors, which would not have been possible with natural images such as the flowers in Fig. 1A. In this sense, the present shapes might be thought of as “Petri dishes” for complexity: Stripped-down environments that allow us to isolate complexity *per se*. Second, there is a long history of studying shapes such as these (indeed, they are often referred to as “Attneave shapes,” for first being used by Attneave & Arnoult, 1956), and yet in this long history there has never been a deeply satisfying account of what makes such shapes appear simple or complex. By exploring such an account here, we hope to validate an approach that could later apply more broadly. Third, and finally, these unusual shapes are deeply unfamiliar, with no pre-existing associations. By contrast, an analysis of the complexity of, say, works of fine visual art would inevitably have to consider more than just the colors and strokes in a painting but also the underlying message of the artist, the emotions the work evokes, and even the political and cultural context in which it was created. All of these high-level factors may influence the subjective experience of complexity, but in ways that could never be captured by analyzing the images themselves. Indeed, given our goal to isolate *visual* complexity as distinct from other notions of complexity, it may well be that art and other such media

simply are not open to this kind of analysis, since they may be complex in ways that go beyond visual perception itself. In other words, unlike the stimuli used in the present work, a painting or artifact's complexity may involve factors that are not truly *seen* at all but rather are genuinely *conceived*, in ways that may be interesting for other purposes (Ahl & Keil, 2017; Kominsky, Zamm, & Keil, 2018) but perhaps not for investigations of visual complexity per se.

Still, we fully embrace the existence of these broader notions of complexity, and we think that information-theoretic approaches such as those explored here do hold promise for at least some of these other complexity-relevant experiences—as in, for example, work in perceptual organization that characterizes visual patterns in terms of the “bits” required to encode them (Chater, 1996; Koenderink, Van Doorn, & Pinna, 2018; Van der Helm, 2000). An array of colors and shapes (of the sort that might appear in a painting) might well be characterized in terms of its predictability (e.g., from one element to the next, or from the process that generated them), just as the shapes explored here were characterized in terms of the predictability of their internal structure. And the same could be true of even broader perceptual features, such as patterns of motion (e.g., whether an object's future trajectory is predictable from its past movements) or even sequences of auditory stimuli (Kidd, Piantadosi, & Aslin, 2014). In this way, it seems to us that the promise of information-theoretic approaches to visual complexity has hardly been fully realized, including even in the present work; indeed, we hope and expect that future work will take such an approach to these broader notions of complexity, where we believe they will again help to shed new light on such experiences.

9.3. *Complexity and attention*

Another outstanding question is the precise mechanism by which complexity guides attention. In Experiments 1–3, we found a search asymmetry in which complex targets were easier to find among simple distractors than vice versa. Does this mean that complexity is a “basic visual feature,” as search asymmetries are often used to indicate?

As noted earlier, there is an active debate in the literature over exactly what constitutes efficient processing and what this implies about the mind more generally (Rauschenberger & Yantis, 2006; Rosenholtz, 2001; Wolfe & Utochkin, 2019). Indeed, one possibility is that efficient search in such displays is driven not by privileged processing of the targets but rather by facilitated rejection of the distractors, especially when they are homogeneous. For example, consider our Experiment 1 (e.g., the schematic display appearing in Fig. 3): Perhaps it is easier to determine that there are many identical objects on the display when most of those objects are simple than when they are complex, which could manifest as faster response times for “complex target” trials (which have more simple distractors) than “simple target” trials (which have more complex distractors). Of course, this is a possibility with any classical search asymmetry (including Treisman's original discoveries) and need not reject notions of privileged processing; this may simply be the mechanism by which certain features guide attention

in visual scenes. However, one reason to suspect that distractor rejection is not the (sole) driver of the present results is a pattern observed here at the lowest set size. In that experiment, it was indeed true that “all simple” trials were easier for subjects than “all complex” trials (which is consistent with it being relatively easier to determine that there are simple distractors than complex ones)—but this was not true at set size 6 (where search is less effortful), where that difference was not present (or, if anything, was reversed). Tellingly, at that set size, subjects were still significantly faster to find complex targets among simple distractors than to find simple targets among complex distractors, even though they did not find it any easier to identify “all simple” trials than “all complex” trials. Put differently, $RT(\text{“simple target”}) - RT(\text{“complex target”})$ was greater than $RT(\text{“all complex”}) - RT(\text{“all simple”})$, $t(11) = 2.28$, $p < .05$, $d = 0.66$. Though we do not wish to read too deeply into this result (and the experiments were not designed with this analysis in mind), it nevertheless suggests that even when searching through complex distractors cost no more time than searching through simple distractors, subjects were still more efficient in finding complex targets among simple distractors than vice versa.⁶

Although the above piece of evidence hints that the efficient search we observed here is guided (at least in part) by the complex targets rather than (merely) by rejecting simple distractors, future studies could address this issue more comprehensively by exploring search displays with heterogeneous distractors. However, given that complexity is a continuous feature (and comes in “degrees”), search asymmetry might be poorly suited to displays with heterogeneous search arrays, because any heterogeneous search array will always have a most complex or simplest target. In that case, a different approach could be to show a “preview” of a given target (simple or complex) at the beginning of each trial and ask subjects to judge if that object appears in the subsequent display of heterogeneous objects. If complexity guides attention efficiently, subjects could be faster to find the target when it is complex than when it is simple, even in heterogeneous arrays.⁷ Another possibility could be to look for complexity-based search facilitation even when complexity is not the property of interest: For example, subjects could be faster to find a target defined by one texture (e.g., a wooden object among paper objects) if the texture singleton is also more complex than if it is simple.

9.4. Complexity’s causal arrow

Experiments 4 and 5 jointly revealed a relationship between the complexity of objects and exploratory behavior. Compared to simple objects, complex objects were associated with prolonged viewing. Though we interpreted this result as showing that perceived complexity *promotes* such viewing, it is possible that the opposite is true: Perhaps more complex stimuli take longer to process (as suggested by Lewis & Frank, 2016), and the mind then forms an impression of complexity on the basis of that prolonged viewing (or some related notion of perceived effort).⁸

Though the data from those experiments cannot speak directly to this possibility, one reason to think that complexity impressions precede (rather than follow) increased

engagement comes from our prior experiments. Experiments 1–3, using visual search, demonstrated that the visual processing required to identify a particular target shape (e.g., a complex target among simple distractors) takes only several hundred milliseconds. By contrast, the exploratory behaviors in Experiments 4–5 often consume several seconds. In that case, assuming that complexity impressions are generated by similar mechanisms in both cases, it seems more parsimonious to conclude that (fast) complexity impressions arise before (slower) prolonged viewing. However, future work that actively manipulates exploration could settle this question more decisively.

9.5. *Minding complexity*

Beyond the theoretical conclusions of the present work, a more practical consequence of these experiments may be their implications for other work that may wish to manipulate or control for complexity. The fact that complexity may be automatically extracted by the visual system (as suggested by our Experiments 1–3) and produce effects on voluntary engagement (as suggested by our Experiments 4–5) suggests that the complexity of a stimulus could be a relevant factor in a wide variety of cognitive tasks or settings. Indeed, the complexity of a stimulus is already thought to play a role in cognitive processes as diverse as mental rotation (Bethell-Fox & Shepard, 1988), word learning (Lewis & Frank, 2016), infant play (Brennan, Ames, & Moore, 1966), working memory (Alvarez & Cavanagh, 2004), and aesthetic beauty (Spehar, Clifford, Newell, & Taylor, 2003; Van Geert & Wagemans, 2020)—and there may well be many more such processes for which complexity could be relevant.

However, all such investigations face the familiar challenge of characterizing complexity in the first place—a challenge so great that some studies have had to first acquire other behavioral evidence of a stimulus's complexity (as in Alvarez & Cavanagh, 2004), use extremely impoverished stimuli (as in Chipman & Mendelson, 1979), or simply make do with measures that fail to separate complexity per se from other confounding factors (perhaps as in Brennan et al., 1966). By contrast, our approach here would allow such work to produce an essentially endless library of diverse visual stimuli and to compute their complexity precisely and in a psychologically plausible way (and without needing to first measure human performance on such stimuli). We thus hope this work will be important not only for understanding the psychological experience of complexity but also for manipulating and measuring such factors in future work.

10. Conclusion

The present work is a case study of how an experience as broad and high-level as complexity can be studied in a psychologically plausible and computationally rigorous framework, and in a way that separates complexity itself from factors that might merely be confounded with it. This approach allowed us not only to measure complexity in an objective way but also to study its nature and function in the mind,

where it appears that complexity is a privileged visual signal for engaging with one's environment.

Acknowledgments

For helpful discussion and/or comments on previous drafts, we thank Kyle Cave, Jacob Feldman, Gary Lupyan, Jonathan Kominsky, and members of the JHU Vision Sciences Group. This work was supported in part by the JHU Science of Learning Institute.

Open Research badges



This article has earned Open Data and Open Materials badges. Data and materials are available at <https://osf.io/aev3j/>.

Notes

1. Here, we used the medial axis transform (MAT) to compute skeletal representations of visual shape. By contrast, Feldman and Singh (2006) used the maximum a posteriori skeleton (MAP), which is a Bayesian estimation of a shape's generative skeleton that penalizes extra branches. Generally, MAP only retains the axes of MAT that correspond to highly salient "parts" of a shape, and removes those branches induced by relatively smaller contour changes. However, for the shapes used in our experiments, the MAP procedure produces counterintuitive skeletons that fail to capture shape properties that seem subjectively important (including not only here but also in previous work; Firestone & Scholl, 2014), and so it may not be well-suited for the kinds of stimuli we explore here. Indeed, if anything MAP-based measure of description length was *negatively* correlated with subjective complexity ratings ($r = -.49$), whereas MAT-based description length was positively correlated (see main text for details). We thank Jacob Feldman for discussion and insight on these issues.
2. Because all trials in this experiment were shown with set size 18, it is natural that these search times were longer than in Experiment 1.
3. We chose this design because a pilot study, which revealed a marginally significant effect in our predicted direction, also caused several subjects to report losing patience and concentration near the end of the experiment. This may be because the SHINE procedure lowers the overall contrast of the shapes it processes such that it was simply more difficult to distinguish each shape from its background (regardless of how difficult it was to distinguish different shapes from each other). An

exploratory reanalysis of those pilot data revealed a powerful search asymmetry in the first half of the session; so, in the present experiments, we ran twice as many subjects in “half” a session’s worth of trials, and pre-registered this design and sample size (see main text for details).

4. Of course, it is not even clear that this is the best strategy: Complex objects may well be more distinguishable from one another than simple objects; and if that is the case, it might be better to spend more time studying simple objects than complex ones.
5. Indeed, if we re-examine the results of Experiment 6 and compare the predictive power of structural surprisal to the power of *number of branches* (a measure of the number of object parts), number of branches is unable to explain any variance that is unexplained by surprisal, whereas surprisal explains variance that is unexplained by number of branches ($\beta = 2.0$, $p < .0001$).
6. We thank a reviewer for prompting this analysis, which suggests that complex targets per se at least partly guide efficient search in this task.
7. However, even *this* experiment may not fully resolve the relevant issue. For example, simple and complex objects load on visual working memory differently Alvarez and Cavanagh (2004), which, in turn, could interfere with subsequent search performance (since the items must be held in memory). But these considerations extend far beyond the present studies of complexity and testify to the inherent difficulty isolating any one feature as an attentional guide in visual search.
8. We thank a reviewer for raising this possibility.

References

- Ahl, R. E., & Keil, F. C. (2017). Diverse effects, complex causes: Children use information about machines’ functional diversity to infer internal complexity. *Child Development*, 88(3), 828–845.
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15(2), 106–111.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61(3), 183–193.
- Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology*, 53(4), 221–227.
- Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, 53(6), 452–471.
- Ayzenberg, V., Chen, Y., Yousif, S. R., & Lourenco, S. F. (2019). Skeletal representations of shape in human vision: Evidence for a pruned medial axis model. *Journal of Vision*, 19, 6.
- Ayzenberg, V., & Lourenco, S. F. (2019). Skeletal descriptions of shape provide unique perceptual information for object recognition. *Scientific Reports*, 9, 9359.
- Berlyne, D. E. (1957). Conflict and information-theory variables as determinants of human perceptual curiosity. *Journal of Experimental Psychology*, 53(6), 399–404.
- Berlyne, D. E. (1958). The influence of complexity and novelty in visual figures on orienting responses. *Journal of Experimental Psychology*, 55(3), 289–296.
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw-Hill.
- Berlyne, D. E. (1966). Curiosity and exploration. *Science*, 153(3731), 25–33.
- Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, 8(5), 279–286.

- Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 12–23.
- Blum, H. (1973). Biological shape and visual science (part i). *Journal of Theoretical Biology*, 38(2), 205–287.
- Brennan, W. M., Ames, E. W., & Moore, R. W. (1966). Age differences in infants' attention to patterns of different complexities. *Science*, 151(3708), 354–356.
- Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review*, 103(3), 566–581.
- Chipman, S. F., & Mendelson, M. J. (1979). Influence of six types of visual structure on complexity judgments in children and adults. *Journal of Experimental Psychology: Human Perception and Performance*, 5(2), 365–378.
- Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, 8(3), e57410.
- Cutting, J. E., & Garvin, J. J. (1987). Fractal curves and complexity. *Perception & Psychophysics*, 42(4), 365–370.
- Day, H. (1968). The importance of symmetry and complexity in the evaluation of complexity, interest and pleasingness. *Psychonomic Science*, 10(10), 339–340.
- De Winter, J., & Wagemans, J. (2006). Segmentation of object outlines into parts: A large-scale integrative study. *Cognition*, 99(3), 275–325.
- Donderi, D. C. (2006). Visual complexity: A review. *Psychological Bulletin*, 132(1), 73–97.
- Feldman, J., & Singh, M. (2005). Information along contours and object boundaries. *Psychological Review*, 112(1), 243–252.
- Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences of the United States of America*, 103(47), 18014–18019.
- Firestone, C., & Scholl, B. J. (2014). “Please tap the shape, anywhere you like”: Shape skeletons in human vision revealed by an exceedingly simple measure. *Psychological Science*, 25(2), 377–386.
- Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C. J., & Sawey, M. (2011). Predicting beauty: Fractal dimension and visual complexity in art. *British Journal of Psychology*, 102(1), 49–70.
- Halberda, J. (2019). Perceptual input is not conceptual content. *Trends in Cognitive Sciences*, 23(8), 636–638.
- Heaps, C., & Handel, S. (1999). Similarity and features of natural textures. *Journal of Experimental Psychology: Human Perception and Performance*, 25(2), 299–320.
- Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision Research*, 45(13), 1707–1724.
- Hulleman, J., Te Winkel, W., & Boselie, F. (2000). Concavities as basic features in visual search: Evidence from search asymmetries. *Perception & Psychophysics*, 62(1), 162–174.
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, 7(5), e36399.
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, 85(5), 1795–1804.
- Koenderink, J., Van Doorn, A., & Pinna, B. (2018). Measures of prägnanz? *Gestalt Theory*, 40(1), 7–28.
- Kominsky, J. F., Zamm, A. P., & Keil, F. C. (2018). Knowing when help is needed: A developing sense of causal complexity. *Cognitive Science*, 42(2), 491–523.
- Lewis, M. L., & Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition*, 153, 182–195.
- Lowet, A. S., Firestone, C., & Scholl, B. J. (2018). Seeing structure: Shape skeletons modulate perceived similarity. *Attention, Perception, & Psychophysics*, 80, 1278–1289.
- Machado, P., Romero, J., Nadal, M., Santos, A., Correia, J., & Carballal, A. (2015). Computerized measures of visual complexity. *Acta Psychologica*, 160, 43–57.
- Madan, C. R., Bayer, J., Gamer, M., Lonsdorf, T. B., & Sommer, T. (2017). Visual complexity and affect: Ratings reflect more than meets the eye. *Frontiers in Psychology*, 8, 2368.

- Marin, M. M., & Leder, H. (2013). Examining complexity across domains: Relating subjective and objective measures of affective environmental scenes, paintings and music. *PLoS ONE*, 8(8), e72412.
- Mavrides, C. M., & Brown, D. (1969). Discrimination and reproduction of patterns: Feature measures and constraint redundancy as predictors. *Perception & Psychophysics*, 6(5), 276–280.
- Oliva, A., Mack, M. L., Shrestha, M., & Peeper, A. (2004). Identifying the perceptual dimensions of visual complexity of scenes. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26), 1041–1046.
- Op de Beeck, H., & Wagemans, J. (2001). Visual object categorisation at distinct levels of abstraction: A new stimulus set. *Perception*, 30(11), 1337–1361.
- Peirce, J. W. (2007). Psychopy—Psychophysics software in python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, 46(28), 4646–4674.
- Rauschenberger, R., & Yantis, S. (2006). Perceptual encoding efficiency in visual search. *Journal of Experimental Psychology: General*, 135(1), 116–131.
- Rosenholtz, R. (2001). Search asymmetries? What search asymmetries? *Perception & Psychophysics*, 63(3), 476–489.
- Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, 7(2), 1–17.
- Schlochtermeyer, L. H., Kuchinke, L., Pehrs, C., Urton, K., Kappelhoff, H., & Jacobs, A. M. (2013). Emotional picture and word processing: An fMRI study on effects of stimulus complexity. *PLoS ONE*, 8(2), e55619.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.
- Siddiqi, K., Tresness, K. J., & Kimia, B. B. (1996). Parts of visual form: Psychophysical aspects. *Perception*, 25(4), 399–424.
- Singh, M., & Hoffman, D. D. (2001). Part-based representations of visual shape and implications for visual cognition. *Advances in Psychology*, 130, 401–459.
- Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, 28(4), 516–536.
- Spehar, B., Clifford, C. W., Newell, B. R., & Taylor, R. P. (2003). Universal aesthetic of fractals. *Computers & Graphics*, 27(5), 813–820.
- Spehar, B., Walker, N., & Taylor, R. P. (2016). Taxonomy of individual variations in aesthetic responses to fractal patterns. *Frontiers in Human Neuroscience*, 10, 350.
- Switzky, H. N., Haywood, H. C., & Isett, R. (1974). Exploration, curiosity, and play in young children: Effects of stimulus complexity. *Developmental Psychology*, 10(3), 321–329.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95(1), 15.
- Van der Helm, P. A. (2000). Simplicity versus likelihood in visual perception: From surprisals to precisals. *Psychological Bulletin*, 126(5), 770–800.
- Van Geert, E., & Wagemans, J. (2020). Order, complexity, and aesthetic preferences for neatly organized compositions. *Psychology of Aesthetics, Creativity, and the Arts*, 14(2), 135–154.
- VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision Research*, 46(18), 3017–3027.
- Wagemans, J., De Winter, J., de Beeck, H. O., Ploeger, A., Beckers, T., & Vanroose, P. (2008). Identification of everyday objects on the basis of silhouette and outline versions. *Perception*, 37(2), 207–244.
- Watson, A. B. (2011). Perimetric complexity of binary digital images: Notes on calculation and relation to visual complexity. *The Mathematica Journal*, 14, 1–41.

- Wilder, J., Feldman, J., & Singh, M. (2011). Superordinate shape classification using natural shape statistics. *Cognition*, *119*(3), 325–340.
- Wilder, J., Feldman, J., & Singh, M. (2016). The role of shape complexity in the detection of closed contours. *Vision Research*, *126*, 220–231.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The shine toolbox. *Behavior Research Methods*, *42*(3), 671–684.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (pp. 13–73). London: University College Press.
- Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics*, *63*(3), 381–389.
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 0058.
- Wolfe, J. M., & Utochkin, I. S. (2019). What is a preattentive feature? *Current Opinion in Psychology*, *29*, 19–26.