Sustained representation of perspectival shape

Jorge Morales[a,1], **Axel Bax**[a], **and Chaz Firestone**[a,b,c,1]

[a]Department of Psychological & Brain Sciences, Johns Hopkins University, Baltimore, MD 21218; [b]Department of Philosophy, Johns Hopkins University, Baltimore, MD 21218; and [c]Department of Cognitive Science, Johns Hopkins University, Baltimore, MD 21218

Edited by Dale Purves, Duke University, Durham, NC, and approved April 20, 2020 (received for review January 14, 2020)

Arguably the most foundational principle in perception research is that our experience of the world goes beyond the retinal image; we perceive the distal environment itself, not the proximal stimulation it causes. Shape may be the paradigm case of such "unconscious inference": When a coin is rotated in depth, we infer the circular object it truly is, discarding the perspectival ellipse projected on our eyes. But is this really the fate of such perspectival shapes? Or does a tilted coin retain an elliptical appearance even when we know it's circular? This question has generated heated debate from Locke and Hume to the present; but whereas extant arguments rely primarily on introspection, this problem is also open to empirical test. If tilted coins bear a representational similarity to elliptical objects, then a circular coin should, when rotated, impair search for a distal ellipse. Here, nine experiments demonstrate that this is so, suggesting that perspectival shapes persist in the mind far longer than traditionally assumed. Subjects saw search arrays of three-dimensional "coins," and simply had to locate a distally elliptical coin. Surprisingly, rotated circular coins slowed search for elliptical targets, even when subjects clearly knew the rotated coins were circular. This pattern arose with static and dynamic cues, couldn't be explained by strategic responding or unfamiliarity, generalized across shape classes, and occurred even with sustained viewing. Finally, these effects extended beyond artificial displays to real-world objects viewed in naturalistic, full-cue conditions. We conclude that objects have a remarkably persistent dual character: their objective shape "out there," and their perspectival shape "from here."

shape | perspective | constancy | representation | philosophy

Look at the golden object in Fig. 1. What shape is it? Though it projects an ellipse on the back of the eye, we can tell that it is really a circle—a "coin" rotated in depth. This experience illustrates the phenomenon of perceptual constancy, a fundamental process by which the mind goes beyond the two-dimensional (2D) images reaching the eyes to experience the three-dimensional (3D) world that gave rise to such images—e.g., to infer a 3D circular disk from the 2D ellipse it projects (1, 2).

But what is the psychological status of the projected "ellipse" from which we infer the coin's 3D shape? Do we "see" the rotated coin as elliptical and then only "decide" or "judge" that it is a circular 3D object? Or is the reverse true, such that we see its 3D shape and then only realize through effortful reflection that it projects an ellipse? Or is there some third option wherein both experiences coexist in the mind—such that the tilted coin looks, in some sense, both elliptical and circular at the same time?

Perception in Perspective

The place of one's own perspective in visual experience is one of the oldest and most foundational questions in the psychology and philosophy of perception, in part because it poses a puzzle: How does objective knowledge of the world arise from our inherently subjective experience of it? Locke, Hume, and other British Empiricists famously held that we perceive only the 2D projective or perspectival properties of the world, and merely understand or judge the 3D environment that gave rise to such properties: "When we set before our eyes a round globe [...] it is certain that the idea thereby imprinted on our mind is of a flat circle, variously shadowed, with several degrees of light and brightness" (ref. 3, para. II, ix, 8). By contrast, scholars such as Helmholtz and Gibson believed that we primarily perceive 3D distal properties, and that we can appreciate perspectival properties only through effort and reflection (as when a painter requires considerable training to capture 3D images on a 2D canvas): "No one ever saw the world as a flat patchwork of colors—no infant, no cataract patient, and not even Bishop Berkeley [...]. The notion of a patchwork of colors comes from the art of painting, not from any unbiased description of visual experience" (ref. 4, p. 286).

Contemporary vision science has largely upheld the latter view, at least insofar as perception, action, attention, and other psychological processes are widely thought to track the distal properties of objects (i.e., the circularity of the object in Fig. 1), rather than their perspectival properties (i.e., the ellipticity of that object's projection). This can be seen both in foundational work in cognitive science (e.g., refs. 5 and 6) and also in popular vision science textbooks. Gregory (7) and Palmer (8), for example, suggest that perspectival properties exist merely as projections on our retinas, and do not persist in mental representations that guide attention or enter conscious experience: "Perception involves a kind of inference from sensory data to object-reality. Further, behavior is not controlled directly by the data, but by the solutions to the perceptual inferences from the data. If I put a book on a table [. . .] I act according to the inferred physical

PSYCHOLOGICAL AND COGNITIVE SCIENCES

Significance

Our experience of the world goes beyond the light reaching our eyes: Even when a coin is tilted in depth, we infer the circular shape it truly has, rather than the ellipse it projects. But do we ever escape the perspective from which we view the world? A centuries-old philosophical debate asks whether objects of different distal shapes bear a representational similarity to one another when their perspectival shapes match. Here, we test this question empirically. We demonstrate such representational similarity by showing that perspectival shapes influence basic mechanisms of perception and attention, even after distal shape is known. Objects are stamped with the perceiver's perspective: We do not see the world completely separate from our point of view.

Author contributions: J.M., A.B., and C.F. designed research; J.M. and A.B. performed research; J.M. analyzed data with input from C.F.; and J.M. and C.F. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: All data and materials for every experiment reported here are available at Open Science Framework, https://osf.io/thj6y/.

[1]To whom correspondence may be addressed. Email: jorge.morales@jhu.edu or chaz@jhu.edu.

First published June 12, 2020.

Fig. 1. A rotated circular coin.

object—table—not according to the brown patch in my eye" (ref. 7, p. 30, see also, ref. 8).[†]

This theoretical approach is reflected in empirical work as well. Studies of the human visual pathway have suggested that "the visual system progressively transforms information from a retinal to an object-centered reference frame whereby retinal size is progressively removed from the representation" (ref. 9, p. 432). Indeed, the primacy of distal representations is so psychologically entrenched as to influence even low-level perceptual phenomena. For example, distal (but not proximal) properties of stimuli take precedence in how we experience simple 2D shapes (10), visual adaptation (11, 12), and afterimages (13, 14). Such representations also extend to phenomena outside visual processing itself, including memory (15), grasping (16), and sketching (17–20). And whereas some theories of object recognition do emphasize so-called "viewpoint-dependent" representations (where objects are thought to be recognized on the basis of specific snapshot-like views), even such models suggest that the ultimate goal of such processes is a representation of the relevant distal objects that abstracts away from one's particular perspective (e.g., ref. 21).

The Return of the Retinal Image

Despite the apparent scientific consensus about the primacy of distal properties in perception, this question is actually far from settled in nearby literatures, including in the philosophy of perception, where the psychological status of perspectival properties is intensely debated (22–38) (for reviews, see refs. 39 and 40). On one hand, some philosophers sympathize with the dominant view in psychology and vision science, arguing that perceptual experiences are primarily or exclusively about distal, environmental properties (22, 23, 34). At the same time, however, many other philosophers dissent, arguing that our visual experiences are better described by a "dual" character, such that perceptual experience reflects both the true distal properties of objects and their perspectival properties—a circle and an ellipse at the "same time," as it were (24–32, 35–38).

Why the disconnect between these literatures? One reason might be philosophers' interest in visual experience itself—how things *look* to an observer—rather than whatever computational processes give rise to object representations. In other words, all parties agree that the mind can "recover" or "infer" the perspectival properties of objects in one way or another (that, after all, must be how artists and graphic designers are able to capture 3D images on 2D surfaces). Instead, the philosophical interest is in whether such 2D perspectival properties are ever truly "perceived"—rather than merely inferred, imagined, considered, traced, judged, and so on. Yet, despite the long history and philosophical centrality of this question (and even suggestions that empirical data could be relevant; refs. 29, 40, and 41), all such philosophical debates—in both their classical and contemporary incarnations—have traditionally involved only appeals to introspection. For example, Schwitzgebel, a defender of the distal-only view, writes of a rotated coin: "[A]s I stare at the penny now, I'm inclined to say it looks just plain circular, in a three-dimensional space—not elliptical at all, in any sense or by any effort I can muster" (ref. 22, p. 590, also ref. 23). In a similar vein, Smith writes: "the suggestion that pennies, for example, look elliptical when seen from most angles is simply not true—they look round" (ref. 34, p. 172).

The Present Experiments: Sustained Representation of Perspectival Shape

Might empirical data address this fundamental and centuries-old deadlock? Related work on perspectival shape and its interaction with perceptual constancy has primarily relied on more qualitative and effortful tasks (including carefully adjusting a shape to match a target, drawing a copy of a shape, or issuing verbal reports after various instructional manipulations; e.g., refs. 17, 18, and 42–46), or has failed to isolate the key issue of interest. By contrast, here we bring a different kind of evidence to bear on this question. We reason that if rotated circular objects (as in Fig. 1) truly exhibit a representational similarity to distally elliptical objects (i.e., objects that are elliptical both in 3D and in 2D), then they should impair visual search for those objects—a faster and more easily quantifiable task that requires little effort or special instruction. In other words, if a subject must locate a distally elliptical object, they should be "distracted" by a rotated circle whose projection matches the shape of their target, in ways that would cause response-time (RT) differences. Alternatively, if our minds do not represent the perspectival ellipticity of a rotated coin and instead represent it only as the circular object it truly is, then a rotated circular coin should not impair search for objectively elliptical objects.

Here, nine experiments address this question, using real-world and computer-based stimuli, static and dynamic depth cues, multiple shape classes, and both speeded and delayed responding. All studies yielded results consistent with the former interpretation, suggesting that the mind represents objects of matching perspectival shape as being perceptually similar—and in ways that bear on enduring philosophical questions about the role of subjectivity in perception.

Experiment 1: Rotated Coins Impair Visual Search for Elliptical Objects

Methods.

Open science practices. All data and materials for every experiment reported here are available on the Open Science Framework, at https://osf.io/thj6y/ (47).

[†]"Perhaps the most fundamental and important fact about our conscious experience of object properties is that they are more closely correlated with the intrinsic properties of the distal stimulus (objects in the environment) than they are with the properties of the proximal stimulus (the image on the retina). This is perhaps so obvious that it is easily overlooked" (ref. 8, p. 312).

Subjects. One hundred subjects were recruited online via Amazon Mechanical-Turk (for validation of this subject pool's reliability, see ref. 48). All subjects in all experiments provided informed consent and were compensated for their contribution (either financially or with course credit). The experiments were approved by the Homewood Institutional Review Board of Johns Hopkins University.

Stimuli. The stimuli in this experiment (and experiments [Exps.] 2 through 7 as well) consisted of highly realistic 1,000 px × 1,000 px images of the sort appearing in Fig. 1, generated using Blender 2.79 (https://www.blender.org/) and rendered using computational resources at the Maryland Advanced Research Computing Center. Stimuli were presented scaled down to 60% to better fit on subjects' displays. Two objects—a target "coin" with an elliptical shape (aspect-ratio 1.41:1), and a distractor coin with a circular shape—appeared atop two pedestals labeled "1" (left pedestal) and "2" (right pedestal) (Fig. 2). The coins were generated with a waffle-like texture on a metallic surface that produced realistic specular and Lambertian reflection, as well as cast shadows. Identical images with empty pedestals were displayed during feedback and intertrial intervals (ITIs).

Procedure. On each trial, subjects saw an image with one target object (a distally elliptical coin) and one distractor object (a distally circular coin). The target elliptical coin was always presented head on, but the distractor circular coin was presented either head on (on half of trials) or rotated 45° about its vertical axis (on the other half of trials). The direction of rotation (clockwise or counterclockwise) as well as the location of the targets and distractors (pedestal 1 or 2) were counterbalanced across trials. Crucially, the objects' sizes and degrees of rotations were chosen so that the perspectival shapes of the elliptical coin and the rotated circular coin were identical (both projecting a similar elliptical shape onto the "camera" that fixed the perspective of the scene). Subjects' task was simply to press whichever key (1 or 2) matched the numeral on the pedestal carrying the truly elliptical coin—a task we expected subjects to perform with near-ceiling accuracy.

Subjects had up to 1 s to respond. If they pressed a key, they received feedback for 500 ms: the coins disappeared from the pedestals and the image's black bounding box turned green (correct) or red (incorrect); the next trial then started automatically after 1 s. If subjects failed to respond, the bounding box turned red and a "too slow!" message appeared above the empty pedestals; subjects could then press the spacebar to begin the next trial. There were six practice trials, followed by 104 experimental trials.

The logic of this design was thus as follows: If the rotated circular coin's shape is genuinely represented by mechanisms of perception and attention as being similar to the shape of the head-on distal ellipse, then it should serve as an effective distractor to a subject looking for a distally elliptical coin, because its matching perspectival shape should compete with the target for the subject's attention. But if the rotated circular coin is seen as just that—a circular coin—then it should not be a very effective distractor for a subject searching for a distal ellipse.

After completing all of the experimental trials, subjects answered a four-question survey in which they saw a single pedestal supporting either 1) a head-on circular coin, 2) a head-on elliptical coin, 3) a circular coin rotated 45° clockwise, or 4) a circular coin rotated 45° counterclockwise (in this order). For each image, they were asked to select one of two radio buttons reading "this is a circular object (viewed head on or at some angle)" or "this is an oval object (viewed head on or at some angle)." (Note that, though we use "ellipse" here in this paper, we used the more familiar [but perhaps less precise] term "oval" in our instructions to subjects.) This survey ensured that subjects appreciated the distal properties of the objects (e.g., to rule out that they misperceived the rotated circular coins as being distally elliptical). Readers can experience this task for themselves at https://perceptionresearch.org/perspective/.

Exclusion criteria. We excluded any subjects who answered fewer than 80% of experimental trials correctly or who gave any incorrect responses in the postexperiment survey. We also excluded trials with incorrect responses or response times below 100 ms. Though we took these steps to ensure high data quality, none of the results we report here depended in any way on these exclusions; in other words, all relevant effects reported below remain statistically significant, in the same direction, even without excluding any subjects or trials.

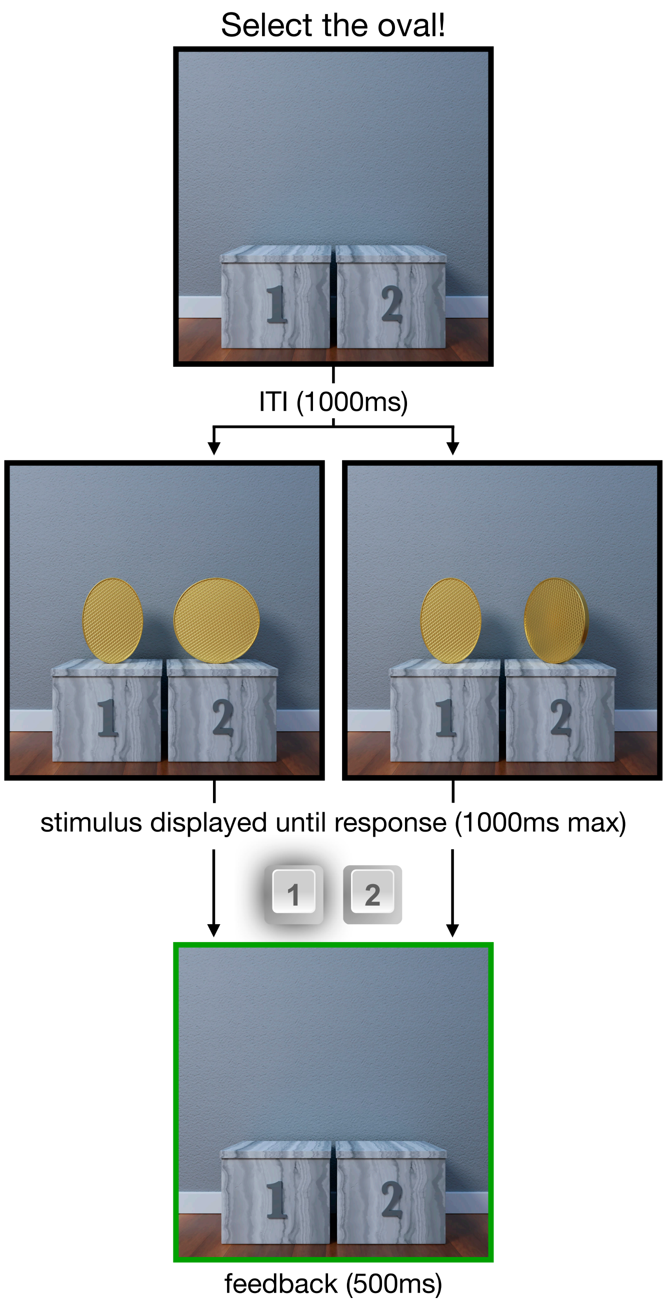

Fig. 2. Task design for experiment 1. Trials started with two empty pedestals labeled 1 and 2, followed by two "coins" appearing atop the pedestals. One of the coins—the target—was always elliptical (here, appearing on pedestal 1). Subjects' task was to press 1 or 2 on their keyboard to indicate the label of the pedestal supporting the target elliptical coin. The other coin—the distractor—was always a circular coin that could appear head on (*Left* stream) or rotated (*Right* stream). After subjects responded, the coins disappeared and visual feedback was provided (green bounding box = correct [depicted]; red bounding box = incorrect or no response). Note that, though we use "ellipse" in this paper, we used the more familiar (but perhaps less precise) term "oval" in our instructions to subjects. The images shown here are brightened for purposes of illustration. Readers can experience the task as subjects did at https://perceptionresearch.org/perspective/.

Results and Discussion. As expected, subjects found this task to be quite easy. Accuracy was 97%, with a mean response time of 518 ms. In other words, subjects identified the truly elliptical coin quickly and accurately.

However, an analysis of response times revealed a striking influence of perspectival shape. Subjects took longer to identify the truly elliptical target object when it was flanked by a rotated circular distractor whose perspectival shape matched the perspectival shape of the target coin (mean RT = 542 ms) than when it was flanked by a distally identical but head-on circular distractor whose perspectival shape differed from the perspectival shape of the target coin (mean RT = 494 ms), $t(63) = 12.48$, $P < 0.0001$ (Fig. 3*A*). This pattern of results was also highly consistent across subjects, with 94% of subjects trending in the expected direction (Fig. 3*B*). In other words, for a subject searching for an elliptical coin, a rotated circular coin (whose perspectival shape

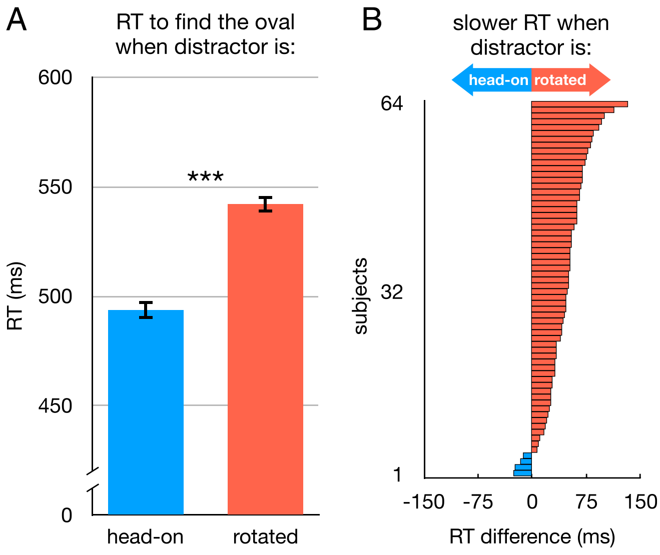

Fig. 3. Results of experiment 1. (*A*) Subjects were slower to select the elliptical coin's position when it was flanked by a rotated circular coin whose perspectival shape matched the target's shape than when it was flanked by the same circular coin seen head on. Bars indicate SEM of the difference between head-on and rotated circular coin trials. ***$P < 0.001$. (*B*) Most subjects (60/64) responded more slowly when the distractor was a rotated circular coin. The graph plots the RT for selecting the elliptical coin flanked by a rotated circular coin minus the RT for selecting it when flanked by a head-on circular coin. Each bar represents one subject.

matched the shape of their target) was a more tempting distractor than a head-on circular coin (whose perspectival shape did not match the shape of their target), despite the fact that the circular distractor had the same distal shape in both cases.

Importantly, there can be no question that subjects understood that the rotated circular object was really a distally circular disk, since success on our postexperiment survey required that subjects correctly identified the true distal shapes of the stimuli. And, again, accuracy on the primary task was near ceiling. This strongly suggests that subjects were not confused at all about the true shapes of the stimuli, and that they were able to discriminate the elliptical target from the circular distractors.

In other words, these results—a robust slowdown when subjects had to find an elliptical coin in the presence of a rotated circular coin—suggest that subjects perceived the elliptical coin and the rotated circular coin as having some property in common (even though they were otherwise completely clear about the different distal shapes of both objects). We thus took this as initial evidence that objects with similar perspectival shapes are represented similarly in the mind, and that these representations are powerful enough to guide attentional processes.

Experiment 2: Controlling for Correlated Properties

We have interpreted the results of experiment 1 as an influence of perspectival shape on perception and attention. However, another possibility is that the observed slowdown instead reflected strategic responding based on some other visual feature, such as area or width. For example, suppose subjects in experiment 1 noticed that the head-on circular coin looked wider (e.g., it had more golden pixels than the elliptical coin); in that case, they may have used this heuristic as an initial response strategy (e.g., "if one of the coins has fewer golden pixels, choose that one") rather than focusing on shape per se.

To rule out this and other similar strategies, experiment 2 repeated the design of experiment 1, but this time the target and distractor coins could have three different sizes: large (roughly 270 px in height, as in experiment 1), medium (165 px), and small (130 px). All possible combinations were shown to subjects during the experiment. For example, a large elliptical coin could appear next to a medium-sized head-on circular coin, a small elliptical coin could appear next to a large rotated circular coin, and so on. Fifty subjects completed 96 experimental trials, with up to 2,000 ms to respond.

This design ensured that response strategies based on area, width, or other such properties would not be effective, encouraging subjects to focus on the objects' shape per se. And indeed, we observed the same interference effect here as in experiment 1, with faster RTs for an elliptical coin beside a head-on circular distractor (685 ms) than for an elliptical coin beside a rotated circular distractor (766 ms), $t\ (24) = 6.15$, $P < 0.0001$ (Fig. 4). This not only served as an effective replication of experiment 1, but also suggested that it was the similarity or dissimilarity in perspectival shape per se (rather than some other factor) that drives these response-time differences.

Experiments 3 and 4: Object Rotation

A different alternative explanation of the present results is just that rotated objects take longer to process than head-on objects, in ways that could produce slower response times on any trials that have rotated objects in them. For instance, objects may be inherently more difficult to recognize when rotated from their canonical views (49, 50); or, subjects might have resorted to mental rotation to normalize the noncanonically presented circular coin into its more familiar head-on presentation (21, 51). In that case, slower responses in trials with a rotated circular distractor could have resulted from the additional time taken to carry out such normalization processes (52, 53), rather than because of matching perspectival shapes.

To rule out these and similar explanations, experiments 3 and 4 repeated the design of experiment 1, but with the addition of different types of rotated distractors. In experiment 3, the two possible distractor objects were a circular coin rotated 45° about its vertical axis (as in the previous two experiments) or a circular coin rotated 75° about the same axis. If rotation per se is causing the slowdown for rotated circular coins, then a 75°-rotated coin should cause a similar or even larger slowdown than the 45°-rotated coin, because it was rotated even further off the frontal plane. But if the interference is caused by matching perspectival shapes, then the 45°-rotated coin should again cause the greatest slowdown (since the 75°-rotated coin did not share the perspectival shape of the head-on elliptical coin).

In experiment 4, the two possible distractor objects were a circular coin rotated 45° about its vertical axis (as in the previous two experiments) or a square object (made of the same material) that was also rotated 45°. Again, if rotation per se is causing the slowdown for rotated circular coins, then a rotated square object should cause a similar slowdown. But if the interference is caused by matching perspectival shapes, then the 45°-rotated coin should again cause the greatest slowdown. Fifty subjects in each experiment completed 96 experimental trials, with up to 1,000 ms to respond.

Both experiments revealed results that were consistent with the matching-perspectival-shapes explanation, but not with the rotation explanation. In experiment 3, subjects displayed faster RTs for selecting an elliptical coin beside a circular coin rotated 75° (519 ms) than for an elliptical coin beside a circular coin rotated 45° (to match the perspectival shape of the target; 550 ms), $t\ (38) = 5.37$, $P < 0.0001$ (Fig. 4). And the same was true in experiment 4: An elliptical coin was easier to find when it appeared next to a rotated square (467 ms) than when it appeared next to a rotated circular coin (530 ms), $t\ (40) = 9.85$, $P < 0.0001$ (Fig. 4).

These results provide further evidence that the interference is caused by the representational similarity produced by similar perspectival shapes per se, in ways that cannot be explained by low-level properties of the rotated circular distractor such as area, width, or rotation itself.

Experiment 5: Dynamic Cues

The stimuli used in the foregoing experiments contained many rich pictorial cues to depth and 3D shape, such as uniform textures that recede in depth according to linear perspective, and realistic lighting and shading. And, indeed, such cues were easily sufficient for subjects to infer the distal 3D properties of the objects, as indicated by answers on the postexperiment surveys. However, being only static images on a display, it is possible that the available cues produced less-than-fully vivid experiences of the objects' distal 3D shapes. Indeed, object recognition in the world often involves dynamic cues to depth and 3D shape, such as motion—whether generated by observers' own movements (e.g., motion parallax) (54) or by the object itself (55). Experiment 5 thus added such dynamic motion cues to the displays, to further emphasize the 3D shape of the objects (and thereby make any influence of perspectival shapes that much more surprising).

Fifty subjects participated in experiment 5, which proceeded identically to experiment 1 except that now the image shown on each trial was not a static frame but instead an animated movie that showed the targets and distractors oscillating back and forth about their vertical axes, in ways that produced a more vivid experience of a 3D object in depth. The head-on objects (i.e., the elliptical target coin and the head-on circular distractor coin) rotated 15° clockwise and counterclockwise from 0°, traversing 30° of total rotation (5° of rotation between each of seven frames). The rotated object (i.e., the rotated circular distractor coin) rotated 6° clockwise and counterclockwise

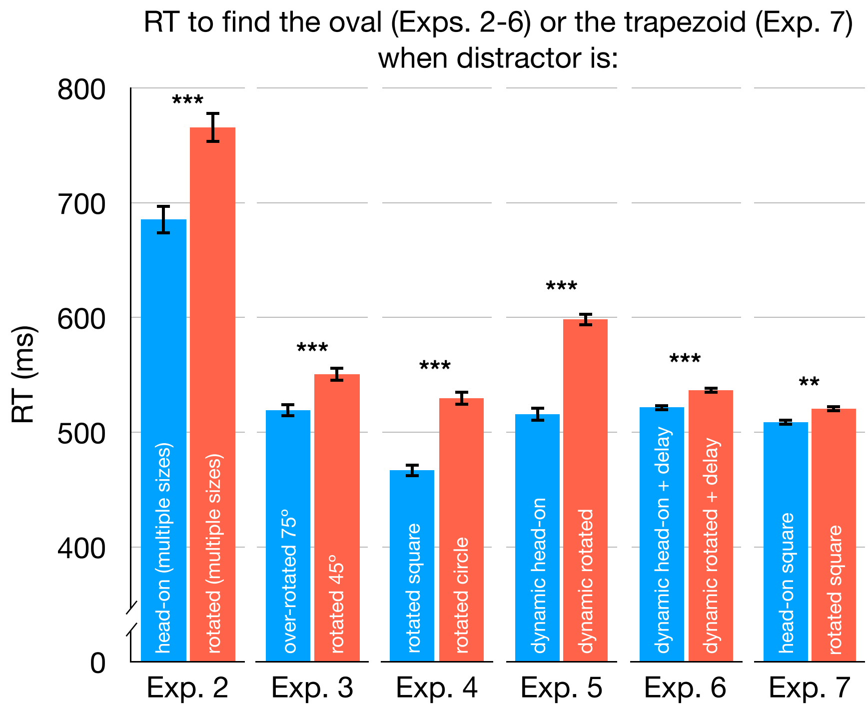

Fig. 4. Results of experiments 2 through 7, all of which revealed a consistent and robust influence of perspectival shapes on perception and attention, even when controlling for confounds that could potentially explain the results from experiment 1. These studies move beyond the idiosyncrasies of the predictivity of object size (Exp. 2), the influence of rotation itself (Exps. 3 and 4), lack of rich dynamic cues (Exp. 5), sufficient time to form a 3D representation (Exp. 6), and any one shape class (Exp. 7). Across these many studies, only the representational similarity between the target and the rotated distractor explained the observed RT slowdown. Bars indicate SEM of the difference between trials with a head-on and a rotated distractor. $^{***}P < 0.001$, $^{**}P < 0.01$.

from 45°, traversing 12° of total rotation (2° of rotation between each of seven frames). These parameters ensured a match in the perspectival shapes of the elliptical coin and rotated circular coin. The coins used during the final survey also contained these dynamic cues.

Even with the addition of these rich, dynamic cues, distractors whose perspectival shapes matched the target shape impaired search performance: 100% of subjects were slower to indicate the location of the target elliptical coin when the distractor was a rotated circular coin (598 ms) than when the distractor was a head-on circular coin (515 ms), t (33) = 13.20, $P < 0.0001$ (Fig. 4). This suggests that the results in previous experiments were not simply driven by the absence of dynamic 3D shape cues, since here the same effects arose even with richer dynamic cues.

Experiment 6: Delayed Responding

A potential concern about all of the preceding experiments is that, even if they do show an influence of matching perspectival shapes on responses, they do not demonstrate any "sustained" representation of such properties. In particular, given that all of the preceding experiments explored only extremely fast responses (with subjects encouraged to indicate the location of the elliptical target coin as soon as they saw it, resulting in response times of ∼450 to 750 ms), one possibility is that perspectival shapes have an influence only on the very earliest stages of visual processing, and only for a very short time. In other words, it might be that the rotated circular coin looks like an ellipse only very briefly, and that this very brief elliptical appearance slows behavioral responses only when those responses are themselves issued very rapidly. If this were the case, it would perhaps undercut the relevance of these findings to the broader theoretical questions under discussion, since those theoretical questions concern more naturalistic (and unrushed) viewing conditions—as when Locke considers what it is like "when we set before our eyes a round globe," or when Schwitzgebel considers his experience "as I stare at the penny now."

To address this worry, experiment 6 proceeded just like experiment 5 (including dynamic depth cues), but this time subjects were prevented from issuing a response until after viewing the coins for a certain period of time. Here, the pedestals on which the coins were sitting began the trial blank (i.e., without the numbers 1 and 2 on them), and remained blank for 1,000 ms even after the coins appeared on top of them. Unlike previous experiments, the numerals did not always appear on the same side; on half of trials, the number 1 appeared on the left pedestal and the number 2 on the right pedestal (as before), but on the other half of trials (randomly interleaved throughout the session), the number 2 appeared on the left and the number 1 on the right. This forced subjects to wait for the labels to appear before they could issue their responses, and so in turn forced them to look at the coins for at least a full second (and often longer). One second is, even by the most conservative estimates, far more than enough time to form a full-fledged 3D representation of an object (56, 57). So, requiring at least this much time to pass ensured that subjects' visual systems would have fully processed the coins' 3D shapes before they could even begin preparing their responses—which in turn ensured that whichever response they did end up giving would reflect a representation of shape that was "complete."

Nevertheless, the same pattern of results was observed: Subjects were slower to indicate the location of the target elliptical coin when the distractor was a rotated circular coin (537 ms after the onset of the numerals, and so 1,537 ms total viewing time) than when the distractor was a head-on circular coin (521 ms after the onset of the numerals, 1,521 ms total viewing time), t (34) = 5.57, $P < 0.0001$ (Fig. 4). Thus, even with more than sufficient time (1 to 2 s) to form a full-fledged 3D representation of an object presented with rich dynamic cues, perspectival shapes still affected perception and attention, in ways that impaired performance on the task.

Note that, in some ways, this result is quite surprising: With sufficient time to resolve the distal properties of the objects, one might not have predicted this sort of sustained interference. One possibility (perhaps to be explored in future work) is that subjects made new eye movements to the objects once the pedestals' numbers were revealed, in ways that "reset" shape constancy computations; subjects may even be more likely to do this when rotated objects are displayed. A different interpretation, though, might be to think of these results as analogous to other more pervasive forms of persistent interference, such as that which occurs in other well-studied phenomena such as the Stroop effect (58, 59) or Garner interference (60).

For example, in the Stroop effect, the word for one color (e.g., "blue") is printed in a different color (e.g., in red ink); subjects who must report the ink's color are slowed by the mismatch between the text's color and its meaning—even though there can be no question that they clearly and unambiguously see the text as red. Thus, these conflicting representations (here, across the visual and semantic domains) continue to compete for the subjects' responses even after subjects have achieved accurate perception of both the words themselves and the colors they are printed in, perhaps in an analogous way to our experiments. So, it may be that rotated circular coins never stop appearing to have something in common with head-on elliptical coins, even after their distal shapes are resolved, and that this induces a kind of persistent response competition.

Perhaps even more related to the present phenomenon is Garner interference. In Garner paradigms, subjects are shown stimuli that may vary along two dimensions (e.g., hue and brightness), and they are asked to discriminate the stimuli based on just one of those dimensions (e.g., to tell the difference between a blue stimulus and purple stimulus, regardless of any differences in brightness). If subjects can discriminate stimuli that vary along only one dimension (e.g., dark blue vs. dark purple) just as easily as they can stimuli that vary along two dimensions (e.g., dark blue vs. light purple), then those two dimensions are taken to be fully separable and processed independently from each other. In contrast, if subjects take longer to discriminate stimuli that vary along both dimensions than stimuli that vary along only one dimension, then these dimensions are considered not to be fully separable. Here, then, distal shapes and perspectival shapes might create a kind of Garner interference: When a given object's distal and perspectival shapes match (e.g., a head-on ellipse that projects an elliptical shape or a head-on circular coin that projects a circular shape), there's no room for these dimensions to interact and, hence, no interference. But when the circular coin is rotated, its distal shape dimension and its perspectival shape dimension differ. The results in experiment 6 with delayed responses are consistent with distal and perspectival shapes being inseparable dimensions in this sense: They cannot be processed in a way that is fully independent from one another [a possibility that has recently been proposed by Lande (31)]. This could potentially explain the persistently slower reaction times in experiment 6 when a rotated coin is used as a distractor: Even after the true distal shapes of the objects are inferred, they cannot be completely separated from their perspectival properties, and so the representational similarity between two objects with shared perspectival shapes continue to interfere with subjects' responses, even after the delay.

Experiment 7: Generalization to Other Shapes

The use of ellipses (of which circles are just a special case) to assess the perception of perspectival shapes has a long history in philosophy and psychology—as when Locke considers perception of a globe, or Kelly and Schwitzgebel consider plates and coins. However, the broader question about representational similarity between objects of matching perspectival shapes concerns any type of shape. In fact, despite their wide use, ellipses have sometimes been considered suboptimal for studying processes related to shape constancy (1), because any distally elliptical object in the world can project any elliptical shape on the retina—which could, in principle, hinder shape constancy processes. Thus, even though subjects in the previous six experiments correctly identified the elliptical target on nearly every trial (and we analyzed reaction time only on those correct trials), it may be possible that the observed response-time slowdown was due to difficulties efficiently achieving shape constancy rather than due to a representational similarity between the perspectival shapes of the elliptical coin and the rotated circular coin.

Importantly, however, these concerns do not apply to quadrilaterals—which, unlike ellipses (or triangles), are very unlikely to produce identical retinal images when their distal shapes are different, and which could perhaps make shape constancy mechanisms more effective (1). Thus, to generalize our results beyond ellipses (and rule out an alternative explanation of our results due to their impoverished shape constancy), experiment 7 replicated the design of the previous experiments, using trapezoids and squares instead of ellipses and circles. The procedure and task design were the same as in experiment 1, but here the target object was a head-on trapezoidal "canvas" consisting of a textured white grid and a dark frame, mounted flush on a wall. This target trapezoid was always presented next to a distractor: a head-on square canvas that hung either flush against the wall (and hence projected a square) or was hinged 45° toward the viewer, thereby projecting the same trapezoidal perspectival shape as the head-on trapezoidal target. One hundred subjects completed 96 experimental trials.

The experimental hypothesis was the same as before: trials in which the subject's trapezoidal target was flanked by a rotated square (whose perspectival shape matched the target) should result in slower responses compared to trials where the same square distractor was seen head on (and, hence, did not share its perspectival shape with the trapezoidal target). And, in fact, exactly this pattern was observed: Subjects displayed slower RTs for a trapezoidal object beside a rotated square distractor (520 ms) than for a trapezoidal object beside a head-on square distractor (509 ms), $t\ (59) = 3.00$, $P < 0.01$ (Fig. 4). This experiment not only replicated the results from the previous experiments, but generalized them across shape classes, demonstrating that these results were not due to any special properties of ellipses.

Experiment 8: Real World Objects under Naturalistic Conditions

The preceding experiments demonstrate sensitivity to perspectival shapes per se, as distinct from strategic responding based on width or area, difficulties processing rotation, an absence of dynamic cues, an artifact of rapid responding, or the idiosyncrasies of a particular shape class. However, all of these experiments share the limitation that they involve only 2D images on computer monitors rather than real 3D objects in the world. On one hand, this computer-graphics approach allows for the creation of minimally paired stimuli (whose background, lighting, and textures were certain to be identical), and fixing precisely properties such as relative size and angle of rotation. However, as images on monitors, they could never be quite as vivid as real objects in the real world—which not only provides an extremely rich environment in terms of cues to depth and 3D shape (including stereopsis), but may also better map onto the scenarios that Locke, Hume, Helmholtz, Gibson, and contemporary theorists have in mind when they consider the experience of looking at a rotated coin.

For these reasons, experiment 8 shared the logic of experiments 1 through 7, but used a modified design that accommodated the use of real-world objects rather than images on computer displays. We custom-manufactured multiple wooden circular and elliptical disks and placed them on two cubby-filled shelves at rotations designed to match (or not match) the disks' perspectival shapes. Importantly, subjects viewed these real-world objects in person, in a well-lit room, under highly naturalistic conditions. If a similar pattern emerges even under these full-cue, real-world viewing conditions, this would provide the strongest evidence yet that perspectival shapes are represented by mechanisms of perception and attention.

Methods.

Subjects. Ten subjects were recruited from Johns Hopkins University. All subjects provided informed consent and received course credit for their participation.
Stimuli. Sixteen wooden "coins"—eight elliptical (major axis: 11.43 cm; minor axis: 8.08 cm) and eight circular (diameter: 11.43 cm)—were laser cut from 0.635-cm-thick oak planks using a Universal Laser Systems VLS 4.60 laser cutter (Fig. 5). The coins were then sanded and varnished with shellac.

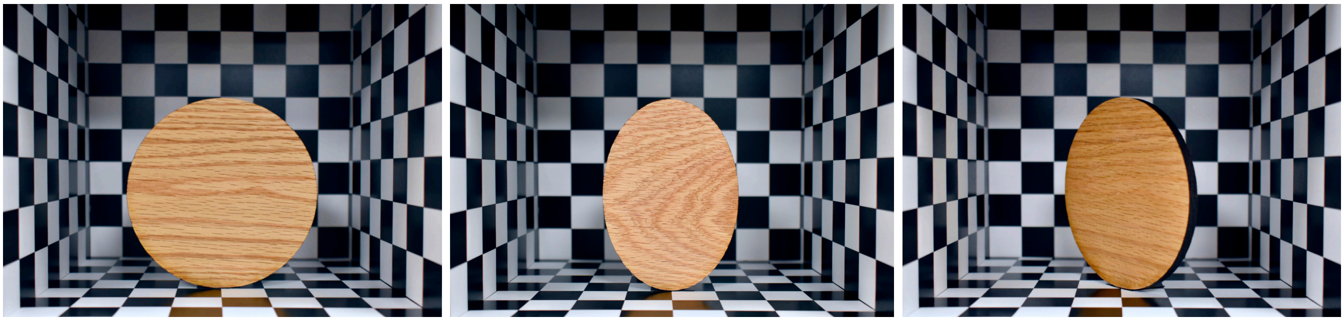

Fig. 5. Example stimuli used in experiments 8 and 9. Circular head-on (*Left*), elliptical (*Center*), and circular rotated (*Right*) laser-cut wooden "coins." Though this figure captures many of the cues available to subjects, readers are reminded that subjects in both of these experiments viewed these objects binocularly (enabling stereopsis) and in naturalistic lighting; and experiment 9 allowed subjects the freedom to move their heads (enabling motion parallax). These factors produced an extremely vivid perception of their 3D shape that a reader cannot fully experience on a screen or a printed page.

Two black shelves (67.3 cm × 67.3 cm × 19.7 cm) with nine cubbies (each 20.7 cm × 20.7 cm × 19.7 cm, arranged in a 3 × 3 grid) were placed on a table (1.5 m × 61 cm × 71 cm) at the end of a large room (7.3 m × 4.5 m) with normal indoor office illumination (Fig. 6*A*). The shelves were parallel to the wall and aligned with each other with a 20.3-cm gap between them. Black cloth boxes covered this gap, and a 2.54-cm white circular plastic disk in the middle was used as an orienting cue. The shelves were 2.5 m away from the far end of another table with two large (6.35 cm in diameter) round response buttons placed on it. The buttons (green and blue Orby switches by P.I. Engineering) were connected via USB to a RaspberryPi 3 computer. Subjects sat behind this table on a chair facing the display, with their heads in a chinrest 33 cm above the surface of the table (such that their eyes were positioned to experience the wooden coins as having their intended perspectival shapes). A head-on wooden coin viewed at eye height thus subtended 2.62° of visual angle. Subjects could view the entire display (i.e., all cubbies in both shelves) without needing to move their heads.

Each shelf contained eight objects (one per cubby, with the central one left empty): four elliptical and four circular coins, each mounted on a thin piece of cardboard (not visible to the subject) that allowed them to stand upright and be rotated as required (Fig. 6*A*). All of the elliptical coins, and two of the circular coins in each shelf, were carefully angled to look head on from the subject's chair. The remaining circular coins were carefully angled to project a similar outline as the elliptical coins from the subject's point of view (i.e., each ellipse and rotated circle projected a similar outline; but depending on which cubby it was placed in, it required a different degree of rotation relative to its cubby). The inner walls of the 16 cubbies containing coins were covered with a black and white checkerboard pattern. Each cubby had two light-emitting diode (LED) lights affixed to the subject-facing end of the upper wall and were connected to a breadboard controlled by the RaspberryPi (this equipment was also not visible to the subject). One LED was placed behind the white circular orienting mark between the shelves. The experiment was programmed in PsychoPy (v.1.83.04) using Python 2.7.13.

Procedure. The subject's goal in the task was the same as in previous experiments—to indicate the location of an elliptical coin—but the task design itself differed (Fig. 6*B*). Each trial began with the central cue flashing for 1,500 ms. This was meant to encourage subjects to begin every trial with their eyes fixed in the middle of the display (though the logic of the experiment did not depend on such fixations). Then, the LEDs of two cubbies, one from each shelf, turned on (and remained on for the duration of the trial). On every trial, one of the cubbies cued by LEDs contained a distally elliptical coin (i.e., the subject's target), whereas the other cued cubby contained a circular coin (i.e., the distractor). On half of trials, the distractor was a head-on circular coin; on the other half of trials, the distractor was a circular coin angled to look 45° rotated on its vertical axis toward the subject. Subjects' task was to select the shelf (left or right) that contained the elliptical coin by pressing the corresponding left or right button on the table. The logic of the design was the same as before: If perspectival shapes impact perception and attention, then it should be harder to indicate the location

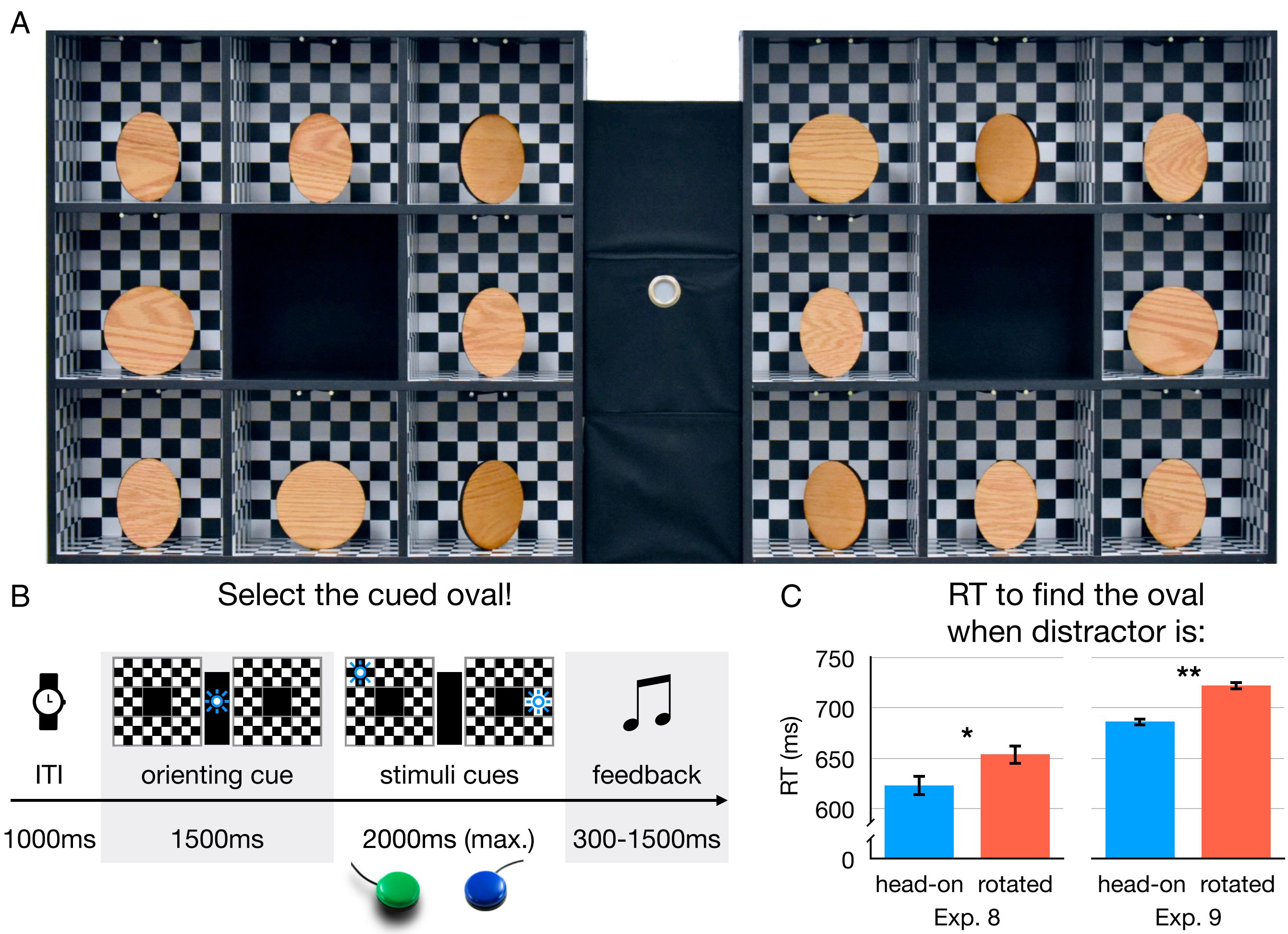

Fig. 6. Design and results of experiments 8 and 9. (*A*) Sample configuration. Two shelves with eight elliptical and circular coins were placed in cubbies with checkerboard patterns. Each cubby had a pair of LEDs, and one LED served as the central orienting cue. (*B*) Task design. After the cue between the two shelves blinked for 1,500 ms, one cubby per shelf was cued with an LED. One of the cued cubbies always contained an elliptical coin and the other one always contained a circular coin (head on and rotated counterbalanced). Subjects pressed one of two buttons to select the shelf with the cued elliptical coin (*Left* or *Right*). Auditory feedback signaled the correctness of their response. (*C*) Results. Subjects were slower to select the shelf with the cued elliptical coin when the other cued shelf contained a circular rotated coin than when it contained a circular head-on coin, regardless of whether their head movements were restricted (Exp. 8) or not (Exp. 9). A video with sample trials appears at https://perceptionresearch.org/perspective/. Bars indicate SEM of the difference between head-on and rotated coin trials. $^{**}P < 0.01$, $^{*}P < 0.05$.

of the elliptical object when its foil is a rotated circular coin (whose perspectival shape matches that of the target) than when its foil is a head-on circular coin (whose perspectival shape differs from that of the target).

The coin configuration in the shelves was determined a priori, and was different for each subject. We created 16 different possible coin configurations, and assigned a given configuration to a given subject pseudorandomly before any data collection began. The configuration remained constant throughout that subject's session. Three constraints were followed when creating these configurations. First, the left and right shelves did not contain the same type of coin in the same position (e.g., if there was an elliptical coin in the upper left cubby of the left shelf, there was a circular coin in the upper left cubby of the right shelf). Second, there was at least one elliptical coin in each column and each row of each shelf. Third, no two circular coins with the same rotation were next to each other. These precautions eliminated certain idiosyncrasies from any given configuration.

Subjects had up to 2,000 ms to respond after the LEDs cued the two coins. They received auditory feedback about their response: If they responded in the allotted time, they heard a short (300 ms), high-pitched sound for correct responses or a short (300 ms), low-pitched sound for incorrect ones. If they failed to respond, they heard a long (1,500 ms), low-pitched sound. In either case, the next trial started automatically, after 1,000 ms. Subjects got acquainted with the task and feedback in 6 practice trials at the beginning of the experiment. These were followed by 192 experimental trials, divided into two blocks with a 30-s break in between. At the end of the experiment, subjects completed a survey administered by the experimenter in which they were asked to identify the true shape of a cued circular head-on coin, an elliptical coin, and a rotated circular coin (in this order) by responding to the question "Is this a circular or an oval object?" Finally, subjects' stereoscopic depth perception was assessed with the circles test from the Randot stereo test (Stereo Optical Company, Inc.), to ensure normal stereopsis.

A detail worth emphasizing about this design and setup is that subjects saw the entire display (including all 16 wooden coins), with both eyes, in a brightly lit room; the coins themselves were also textured and casting shadows; and they appeared against a high-contrast textured background. The conditions were thus ideal for experiencing them as 3D objects rather than as 2D images. If the same interference effect still occurs under these circumstances, it would be especially strong and especially relevant evidence concerning the perspectival appearance of these objects. A video with sample trials appears at https://perceptionresearch.org/perspective/.

Results and Discussion. Even under normal viewing conditions and after a long exposure to maximally naturalistic stimuli, the pattern of results observed in the previous computer-based experiments was observed here in a real-world context. Subjects were slower to select the shelf with the cued elliptical coin when a rotated circular coin was cued on the other shelf (655 ms) compared to when a head-on circular coin was cued (624 ms), t (9) = 2.89, P = 0.018 (Fig. 6*C*). Moreover, 9 subjects' RTs trended in the expected direction; performance was at ceiling (mean correct responses = 96%); all subjects correctly identified the true distal shape of the coins during the postexperiment survey; and all subjects had normal stereoscopic depth perception as revealed by the Randot stereo test (in which they correctly identified the test circle with a disparity of 70 arcsecs or less; median = 40 arcsecs). The effect also persisted throughout the experiment; for example, the mean RT difference between trials with head-on and rotated distractors in the first block of trials was 26 ms, and in the second block it was 35 ms (which, if anything, trends in favor of the effect increasing, not decreasing). This suggests that the interference created by matching perspectival shapes was impervious to practice and learning effects, and thus extremely robust, persistent, and resistant to top-down control (since knowing the objects' true shape failed to eliminate the relevant effects) (61).

Thus, this experiment not only replicates the logic and design of the previous seven experiments, but shows that such effects persist with real-world 3D objects, seen under normal viewing conditions, by subjects with normal vision.

Experiment 9: Real World Objects with No Constraints

Though the previous experiment tested subjects under extremely naturalistic conditions—using real 3D objects appearing in a fully lit room under binocular viewing—they also included a head restraint so as to keep fixed the perspective of subjects (and the perspectival shapes of the circular and elliptical coins). Though this restraint ensured a high degree of precision with respect to perspectival shape, it arguably prevented motion parallax cues, since the subject's head was not free to move. To even further enhance the naturalistic qualities of this experience, experiment 9 replicated experiment 8, holding constant every procedural detail except one: This time, subjects sat in the chair completely unrestrained, such that they were able (if they wished) to move their heads in ways that would generate motion parallax signals. This setup even more closely resembled the conditions described by classical and contemporary theorists—since the experience simply involved looking at tilted objects on a shelf, just as one might outside an experimental context. And so if perspectival shapes interfere here too, this would provide extremely strong and ecologically valid evidence that perspectival shapes are represented by mechanisms of perception and attention.

In fact, subjects were slower to select the shelf with the cued elliptical coin when a rotated circular coin was cued on the other shelf (723 ms) compared to when a head-on circular coin was cued (687 ms), t (9) = 4.03, P = 0.003. This result thus not only replicated the findings from experiment 8, but did so in extremely—even maximally—naturalistic conditions.

Moreover, by providing sufficient time to process the scene, experiments 6 (delayed response), and experiments 8 and 9 (real-world conditions) also allow us to rule out another alternative interpretation of our earlier results. A possible explanation of the results from our earlier experiments (1 through 5, and 7) was a kind of competition for very early visual selection. For example, when searching for a distally elliptical object, the visual system could prioritize processing of elliptical retinal images before other parts of a scene. If this were the case, elliptical projections may attract more processing resources than projections that are unlikely to be caused by distally elliptical objects; in that case, slower RTs could arise in scenes with a rotated coin not because the rotated coin retains an elliptical perspectival appearance in any sustained way, but rather because of competition for very early processes of selection and shape processing. But, in experiments 6, 8, and 9, subjects had plenty of time (from 1 s up to several minutes, rather than milliseconds) to fully process the scene in front of them before making their decision. Thus, even if there was some competition only at the very earliest stages of visual processing, that competition should have been resolved by the time of the subject's response. Yet, we still observed the characteristic pattern of shape interference even under these prolonged-viewing conditions, suggesting that perspectival properties genuinely persist in the mind.[‡] In this way, these results are perhaps consistent with recent views in the philosophy of perception (e.g., ref. 31) holding that perception has an intrinsic "perspectival character" that simply never diminishes or falls away, and so may well predict the kind of lasting interdependence between distal and perspectival contents that we observe here.

General Discussion

Does a rotated coin look elliptical? Though this question has been asked for centuries—since at least Locke and Hume—the nine experiments reported here bring a different kind of evidence to bear on it. Elliptical objects were harder to locate in displays containing rotated circular disks than in displays containing head-on circular disks (experiment 1). These results persisted even when controlling for various low-level confounds—including size, width, and area (experiment 2), as well as rotation itself (experiments 3 and 4)—and also occurred with dynamic depth cues (experiment 5), extra time to generate a stable 3D representation (experiment 6), and other shape classes (experiment 7). Finally, the same pattern generalized from computer-graphics displays to the real world itself, occurring in maximally naturalistic viewing conditions with objects present for several minutes (experiments 8 and 9).

The explanation of these results seems clear and straightforward: An elliptical coin is harder to distinguish from a rotated circular coin (vs. a head-on circular coin) because the two objects *appear to have something in common*. More precisely, when subjects see the rotated circular coin and the head-on elliptical coin, it can be said that they bear a representational similarity to one another (31).[§]

[‡]Yet another, perhaps even more subtle interpretation of these experiments, is that perception of the rotated coin switches back and forth between its distal and perspectival shapes—rather than both interpretations coexisting simultaneously—almost like a bistable figure. [This is roughly the view favored by Kelly (29).] Our findings cannot confirm or disconfirm this interpretation relative to the "co-existing" interpretation. However, we note that both of these views (i.e., coexisting distal and perspectival interpretations vs. switching distal and perspectival interpretations) involve a commitment to perspectival shapes not being discarded, and so both involve a considerably expanded role for perspectival shape than is traditionally favored.

[§]To be even more precise, the results here indicate such representational similarity even without specifying the dimension of such similarity, or the specific features that ground this similarity. For many philosophical issues at stake here, it may be important to distinguish between interference caused by matching perspectival shapes vs. by persisting retinal images themselves vs. by independent representations of ellipticity (31). Our results here cannot adjudicate between these extremely subtle options; but all imply some notion of representational similarity, which is what we take our results to demonstrate.

And so these results suggest that objects' perspectival shapes not only exist as images projected onto our eyes but also guide mechanisms of perception and attention, even once subjects have inferred the true distal shape of objects in their environment.

One aspect of these results that bears repeating—and that distinguishes them from previous work on perspectival shape processing—is that they show the mutual interaction of distal and perspectival shape representations using both highly precise virtual stimuli (i.e., our computer-generated objects) and maximally naturalistic objects (i.e., our laser-cut wooden coins) presented under full-cue conditions. Previous work that explored neighboring questions about shape constancy, perspective, and object recognition used stimuli that were impoverished in at least one of those ways—including work using simple 2D line drawings (10, 15, 17, 18, 51), 2D "shapes" drawn in white noise on a plain background (12), random-dot stereograms (62), objects presented monocularly (45), impoverished depth cues (45), afterimages (12–14), illusions (9, 13, 63), photographs (54, 64, 65), simple and abstract virtual 3D objects (9, 55, 66), and real-world 3D wire objects (67). Importantly, all of these studies—including those using real-world 3D objects (e.g., ref. 54)—presented stimuli at durations lasting from a few milliseconds up to a few seconds at most. In contrast, our stimuli were rich, highly realistic, and in experiments 8 and 9 were presented for several consecutive minutes (since the wooden coins were visible to subjects for the entire experimental session). While the benefits of limiting presentation duration, or using simple stimuli, are clear and perhaps even necessary for certain questions in perception research, the question that motivated Locke, Hume, and more contemporary philosophers and cognitive scientists is primarily concerned with how we see "normal" objects, under normal circumstances, for extended durations that permit close perceptual scrutiny. The results from our computer-based studies demonstrated representational similarity under more typical experimental conditions, whereas our studies with physical objects demonstrated that such representational similarity does in fact extend to normal viewing conditions in the real world.

Moreover, our studies achieve these results using indirect, performance-based, and easily quantified measures derived from a highly naturalistic behavior—locating an object—rather than relying on more subjective, effortful, or instructionally sensitive tasks (such as drawing copies of objects, adjusting standards to match experimental stimuli, and so on). Visual search is, of course, a relatively basic and fast process, but nevertheless it operates over sophisticated and full-blown object representations (rather than just low-level image features). For instance, under normal circumstances, visual search incorporates information about amodal completion, such that search for a circle will be slowed by distractors that are merely amodally completed circles (e.g., a "Pac-Man" shape abutting a square, such that the square appears to be an occluding surface) (68). In that case, visual search operates over the amodally completed object rather than its low-level features. Furthermore, even in the simplest paradigms where targets and distractors differ along just one feature (as ours do in only shape), visual search recruits mechanisms of attentional selection (69) that help disambiguate representationally similar objects (70). Thus, the attentional set brought to bear during visual search makes higher-level representational features of objects available to mechanisms of selection for action and awareness (70–72). These characteristics of visual search suggest that the interference we observed here reflects more than just similarity at very "early" stages of visual processing (which would perhaps be less surprising), and instead that perspectival shapes appear also at higher levels of visual processing where attention, selection for action, and visual awareness are likely to be involved.

What do these results imply? On one hand, there may be something about these results that seems "obvious"; after all, the rotated and elliptical coins project similar shapes on subjects' retinas—why not think that this similarity would interfere with other tasks? But on the other hand, this conclusion is surprising indeed from the orthodox view that perception tracks distal object properties rather than proximal ones. According to this mainstream view, the retinal image (or a representation thereof) is progressively transformed "to an object-centered reference frame whereby retinal (information) is progressively removed from the representation" (9), and it is this object-centered representation that guides perception, attention, action, and so on: "I act according to the inferred physical object [...] not according to the brown patch in my eye" (7). If perspectival information were indeed truly absent from the object representations that inform perception and action, then subjects should not be moved by the coin's perspectival shape. And so the present results are inconsistent with claims that we perceive distal properties only, or that there is nothing shared by two objects with differing distal shapes but matching perspectival shapes. Moreover, the present results are not only inconsistent with certain claims about object-centered reference frames in perception—they are also inconsistent with introspective reports often found in the philosophical literature, such as those claiming that a rotated circular coin "looks just plain circular, in a three-dimensional space—not elliptical at all, in any sense or by any effort" (22). If the contents of perception included only the distal shapes of objects, and never their perspectival shapes, then a circular coin (head on, or not) should not have any shared perceptual contents with an elliptical coin. And so slowed search for an elliptical coin flanked by a rotated circular coin should not occur on this view either.

Instead, and contrary to these positions in the philosophy, psychology, and neuroscience of perception, the present results suggest that we do not see the distal properties of objects in ways that completely separate them from our point of view. Whereas a rotated circular coin and an elliptical coin do not in fact have the same shape in the world, they do share some property in common from our perspective, and our results show that the mind is sensitive to that similarity. Perspectival shapes are sufficiently present in perception to interfere with subjects' search behavior and attentional processing; but at the same time, they are sufficiently distinct from the distal contents of our shape representations that they do not confuse subjects about the true properties of objects. In other words, the sense in which the coin "appears elliptical" is not somehow false or illusory, since 1) the rotated coin truly does subtend a solid angle of the same shape that a frontal ellipse subtends, and is seen that way; and 2) subjects who have this experience of the rotated circular coin do not then perceive the presence of a distally elliptical object in the world. (After all, much of our knowledge of distally circular objects does come from encountering their often-elliptical projections, which the visual system must exploit to accurately perceive such objects) (73, 74).

Taken together, these experiments address a foundational and even philosophical question about the place of one's perspective in perception. But beyond this, they also show how certain questions of this sort are open to empirical test (or at least can be sensitive to empirical data). Just as a proper reading of the relevant philosophical literature revealed that such questions have not been answered by contemporary vision science (indeed, not even by work that specifically explores viewpoint-dependent models of object recognition), a properly aimed set of experiments really *can* address these questions that were previously only amenable to argument based on introspection. We thus believe this work not only offers an answer to such questions, but also serves as a case study of how empirical data can address philosophical questions about the nature of perception itself.

ACKNOWLEDGMENTS. For helpful discussion, technical support, and/or comments on previous drafts, we acknowledge Jimena and Olivia Monjarás, Kate Storrs, and members of the Johns Hopkins University (JHU) Perception & Mind Lab. J.M. was supported by the JHU Office of the Provost. C.F. was supported by the JHU Science of Learning Institute.

1. Z. Pizlo, *3D Shape: Its Unique Place in Visual Perception*, (MIT Press, 2008).
2. Z. Pizlo, Perception viewed as an inverse problem. *Vision Res.* **41**, 3145–3161 (2001).
3. J. Locke, *An Essay Concerning Human Understanding*, (Clarendon Press, 1975).
4. J. J. Gibson, *The Ecological Approach to Visual Perception*, (Houghton Mifflin, 1979).
5. D. Marr, Early processing of visual information. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **275**, 483–519 (1976).
6. D. Marr, H. K. Nishihara, Representation and recognition of the spatial organization of three-dimensional shapes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **200**, 269–294 (1978).
7. R. L. Gregory, *The Intelligent Eye*, (McGraw-Hill, 1970).
8. S. E. Palmer, *Vision Science*, (MIT Press, 1999).
9. S. O. Murray, H. Boyaci, D. Kersten, The representation of perceived angular size in human primary visual cortex. *Nat. Neurosci.* **9**, 429–434 (2006).
10. I. Rock, C. M. Linnett, Is a perceived shape based on its retinal image? *Perception* **22**, 61–76 (1993).
11. D. H. Arnold, A. Birt, T. S. A. Wallis, Perceived size and spatial coding. *J. Neurosci.* **28**, 5954–5958 (2008).
12. K. R. Storrs, D. H. Arnold, Shape aftereffects reflect shape constancy operations: Appearance matters. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 616–622 (2013).
13. I. Sperandio, A. Lak, M. A. Goodale, Afterimage size is modulated by size-contrast illusions. *J. Vis.* **12**, 1–10 (2012).
14. I. Sperandio, P. A. Chouinard, M. A. Goodale, Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nat. Neurosci.* **15**, 540–542 (2012).
15. I. Rock, The orientation of forms on the retina and in the environment. *Am. J. Psychol.* **69**, 513–528 (1956).
16. V. H. Franz, Action does not resist visual illusions. *Trends Cogn. Sci.* **5**, 457–459 (2001).
17. R. H. Thouless, Phenomenal regression to the real object. I. *Br. J. Psychol. Gen. Sect.* **21**, 339–359 (1931).
18. R. H. Thouless, Phenomenal regression to the "Real" object. II. *Br. J. Psychol. Gen. Sect.* **22**, 1–30 (1931).
19. F. Perdreau, P. Cavanagh, Do artists see their retinas? *Front. Hum. Neurosci.* **5**, 171 (2011).
20. F. Perdreau, P. Cavanagh, Is artists' perception more veridical? *Front. Neurosci.* **7**, 155–167 (2013).
21. M. J. Tarr, Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychon. Bull. Rev.* **2**, 55–82 (1995).
22. E. Schwitzgebel, Do things look flat? *Philos. Phenomenol. Res.* **72**, 589–599 (2006).
23. R. E. Briscoe, Vision, action, and make-perceive. *Mind Lang.* **23**, 457–497 (2008).
24. C. Peacocke, *Sense and Content*, (Oxford University Press, 1983).
25. M. Tye, Perceptual experience is a many-layered thing. *Philos. Issues* **7**, 117–126 (1996).
26. A. Noë, *Action in Perception*, (MIT Press, 2004).
27. A. Noë, *Varieties of Presence*, (Harvard University Press, 2012).
28. J. Cohen, Perception and computation. *Philos. Issues* **20**, 96–124 (2010).
29. S. D. Kelly, Content and constancy: Phenomenology, psychology, and the content of perception. *Philos. Phenomenol. Res.* **76**, 682–690 (2008).
30. S. Schellenberg, The situation-dependency of perception. *J. Philos.* **105**, 55–84 (2008).
31. K. J. Lande, The perspectival character of perception. *J. Philos.* **115**, 187–214 (2018).
32. B. Brogaard, Strong representationalism and centered content. *Philos. Stud.* **151**, 373–392 (2009).
33. W. T. Wojtach, Reconsidering perceptual content. *Philos. Sci.* **76**, 22–43 (2009).
34. A. D. Smith, *The Problem of Perception*, (Harvard University Press, 2002).
35. D. Bennett, Seeing shape: Shape appearances and shape constancy. *Br. J. Philos. Sci.* **63**, 487–518 (2012).
36. C. S. Hill, D. Bennett, The perception of size and shape. *Philos. Issues* **18**, 294–315 (2008).
37. D. J. Bennett, The role of spatial appearances in achieving spatial-geometric perceptual constancy. *Philos. Top.* **44**, 1–44 (2016).
38. C. S. Hill, Perceptual relativity. *Philos. Top.* **44**, 179–200 (2016).
39. E. J. Green, S. Schellenberg, Spatial perception: The perspectival aspect of perception. *Philos. Compass* **13**, e12472 (2018).
40. J. Schwenkler, A. Weksler, Are perspectival shapes seen or imagined? An experimental approach. *Phenomenol. Cogn. Sci.* **18**, 855–877 (2019).
41. A. Weksler, Retinal images and object files: Towards empirically evaluating philosophical accounts of visual perspective. *Rev. Phil. Psychol.* **7**, 91–103 (2016).
42. A. H. Sedgwick, "Space perception" in *Handbook of Perception and Human Performance*, K. R. Boff, L. Kaufman, J. Thomas, Eds. (John Wiley and Sons, 1986).
43. A. A. Landauer, Influence of instructions on judgments of unfamiliar shapes. *J. Exp. Psychol.* **79**, 129–132 (1969).
44. A. S. Gilinsky, The effect of attitude upon the perception of size. *Am. J. Psychol.* **68**, 173–192 (1955).
45. W. C. Gogel, The sensing of retinal size. *Vision Res.* **9**, 1079–1094 (1969).
46. V. R. Carlson, Size-constancy judgments and perceptual compromise. *J. Exp. Psychol.* **63**, 68–73 (1962).
47. J. Morales, A. Bax, C. Firestone, Sustained representation of perspectival shape. Open Science Framework. https://osf.io/thj6y. Deposited 3 January 2020.
48. M. J. C. Crump, J. V. McDonnell, T. M. Gureckis, Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS One* **8**, e57410–e57418 (2013).
49. P. Jolicoeur, The time to name disoriented natural objects. *Mem. Cognit.* **13**, 289–303 (1985).
50. R. Lawson, P. Jolicoeur, Recognition thresholds for plane-rotated pictures of familiar objects. *Acta Psychol. (Amst.)* **112**, 17–41 (2003).
51. M. J. Tarr, S. Pinker, Mental rotation and orientation-dependence in shape recognition. *Cognit. Psychol.* **21**, 233–282 (1989).
52. R. N. Shepard, J. Metzler, Mental rotation of three-dimensional objects. *Science* **171**, 701–703 (1971).
53. L. A. Cooper, R. N. Shepard, "Chronometric studies of the rotation of mental images" in *Visual Information Processing*, W. G. Chase, Ed. (Academic Press, 1973), pp. 75–176.
54. D. J. Simons, R. F. Wang, D. Roddenberry, Object recognition is mediated by extraretinal information. *Percept. Psychophys.* **64**, 521–530 (2002).
55. W. Teramoto, B. E. Riecke, Dynamic visual information facilitates object recognition from novel viewpoints. *J. Vis.* **10**, 11 (2010).
56. H. Leibowitz, L. E. Bourne Jr., Time and intensity as determiners of perceived shape. *J. Exp. Psychol.* **51**, 277–281 (1956).
57. M. Bar, A cortical mechanism for triggering top-down facilitation in visual object recognition. *J. Cogn. Neurosci.* **15**, 600–609 (2003).
58. J. R. Stroop, Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **18**, 643 (1935).
59. C. M. MacLeod, Half a century of research on the Stroop effect: An integrative review. *Psychol. Bull.* **109**, 163–203 (1991).
60. W. R. Garner, "The stimulus in information processing" in *Sensation and Measurement: Papers in Honor of S. S. Stevens*, H. R. Moskowitz, B. Scharf, J. C. Stevens, Eds. (Springer Netherlands, 1974), pp. 77–90.
61. C. Firestone, B. J. Scholl, Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behav. Brain Sci.* **39**, e229 (2016).
62. R. A. Champion, D. R. Simmons, P. Mamassian, The influence of object size and surface shape on shape constancy from stereo. *Perception* **33**, 237–247 (2004).
63. D. S. Schwarzkopf, C. Song, G. Rees, The surface area of human V1 predicts the subjective experience of object size. *Nat. Neurosci.* **14**, 28–30 (2011).
64. R. Ellis, D. A. Allport, G. W. Humphreys, J. Collis, Varieties of object constancy. *Q. J. Exp. Psychol. A* **41**, 775–796 (1989).
65. G. W. Humphreys, M. J. Riddoch, Routes to object constancy: Implications from neurological impairments of object constancy. *Q. J. Exp. Psychol. A* **36**, 385–415 (1984).
66. Y. Yamane, E. T. Carlson, K. C. Bowman, Z. Wang, C. E. Connor, A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat. Neurosci.* **11**, 1352–1360 (2008).
67. I. Rock, J. DiVita, A case of viewer-centered object perception. *Cognit. Psychol.* **19**, 280–293 (1987).
68. R. Rauschenberger, S. Yantis, Masking unveils pre-amodal completion representation in visual search. *Nature* **410**, 369–372 (2001).
69. J. S. Joseph, M. M. Chun, K. Nakayama, Attentional requirements in a "preattentive" feature search task. *Nature* **387**, 805–807 (1997).
70. S. J. Luck, M. A. Ford, On the role of selective attention in visual perception. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 825–830 (1998).
71. A. Raffone, N. Srinivasan, C. van Leeuwen, Perceptual awareness and its neural basis: Bridging experimental and theoretical paradigms. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130203 (2014).
72. W. Wu, *Attention*, (Routledge, 2014).
73. D. Purves, B. B. Monson, J. Sundararajan, W. T. Wojtach, How biological vision succeeds in the physical world. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4750–4755 (2014).
74. D. Purves, W. T. Wojtach, R. B. Lotto, Understanding vision in wholly empirical terms. *Proc. Natl. Acad. Sci. U.S.A.* **108** (suppl. 3):), 15588–15595 (2011).