Empirical evidence for perspectival similarity: Comment on Burge & Burge (2022)*

Jorge Morales\textsuperscript{a,b}, Chaz Firestone\textsuperscript{c,d}

\textsuperscript{a}Department of Psychology, Northeastern University  
\textsuperscript{b}Department of Philosophy, Northeastern University  
\textsuperscript{c}Department of Psychological & Brain Sciences, Johns Hopkins University  
\textsuperscript{d}Department of Philosophy, Johns Hopkins University

Abstract

When a circular coin is rotated in depth, is there any sense in which it comes to resemble an ellipse? While this question is at the center of a rich and divided philosophical tradition (with some scholars answering affirmatively and some negatively), Morales et al. (2020, 2021) took an empirical approach, reporting 10 experiments whose results favor such perspectival similarity. Recently, Burge and Burge (2022) offered a vigorous critique of this work, objecting to its approach and conclusions on both philosophical and empirical grounds. Here, we answer these objections on both fronts. We show that Burge and Burge’s critique rests on misunderstandings of Morales et al.’s claims; of the relation between the data and conclusions; and of the philosophical context in which the work appears. Specifically, Burge and Burge attribute to us a much stronger (and stranger) view than we hold, involving the introduction of “a new entity” located “in some intermediate position(s) between the distal shape and the retinal image”. We do not hold this view. Indeed, once properly understood, most of Burge and Burge’s objections favor Morales et al.’s claims rather than oppose them. Finally, we discuss several questions that remain unanswered, and reflect on a productive path forward on these issues of foundational scientific and philosophical interest.

Email addresses: j.morales@northeastern.edu (Jorge Morales), chaz@jhu.edu (Chaz Firestone)

* This is a draft of a response to Burge and Burge (2022), which was accepted at Psychological Review in January 2022 and first shared with us in February 2022. It is available at https://psyarxiv.com/svq39/. Any comments on this document are welcome.

Manuscript draft  
June 1, 2022
Figure 1 contains three images of wooden “coins”: A circular coin viewed head-on, an elliptical coin viewed head-on, and a circular coin rotated at an angle. Consider the following question: Is there any relevant sense in which the rotated circular coin “looks elliptical”? More precisely: Does the circular coin, when rotated, share any aspect of its appearance with the elliptical coin — an aspect it does not share when viewed head-on?

Figure 1: Three wooden “coins”. A head-on circular coin, a head-on elliptical coin, and a rotated circular coin.

This seemingly innocent question is at the center of a rich and divided philosophical tradition, tracing at least to the 17th century (if not further; see Burnyeat 1979) and continuing to the present day. Among the many answers on offer, one class of views embraces *perspectival similarity* (Lande, 2018), holding that the rotated circle and head-on ellipse do indeed share some aspect of their appearance not shared by the head-on circle. Contemporary philosophers who defend perspectival similarity include Peacocke (1983), Noë (2004), Schellenberg (2008), Cohen (2010), Lande (2018), and others. Though these and other theorists differ on the nature of this similarity — i.e., how and why it is that the rotated circle and head-on ellipse look similar, or what grounds this similarity — they accept perspectival similarity in some form, agreeing that there is some relevant sense in which the rotated circular coin looks elliptical.

However, another class of views rejects perspectival similarity. For example, Schwitzgebel (2006) writes of a rotated coin: “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”. Similarly, Smith (2002) writes: “the suggestion that pennies, for example, look elliptical when seen from most angles is simply not true — they look round”.

2 Other theorists who reject perspectival similarity (or at least come very close to doing so) include Briscoe (2008), Hopp (2013), and Siewert (2006). For example, Hopp holds that, when perceived accurately, “round objects tilted away do not look elliptical”.

3 For example, one can deny perspectival similarity while still embracing *perspectival variation*. This view would accept that a rotated circular coin looks different from a head-on circular coin, while still denying that there is any sense in which the rotated circular coin looks elliptical.

2

3
this view hold that there is no sense in which a rotated coin looks elliptical — i.e., no sense in which a circular coin, when rotated, comes to share some aspect of its appearance with a head-on ellipse. 4

Saliently, most of these discussions about shape appearance and visual perspective tend to keep some distance from empirical data. For example, neither Peacocke and Schellenberg (in embracing perspectival similarity) nor Schwitzgebel and Hopp (in rejecting perspectival similarity) discuss or cite any experimental results on shape perception. It is hard to blame them: Though there are of course mountains of empirical work on the perception of 3D shape, there has not (to our knowledge) been an empirical study, using the tools of vision science, directly aimed at this philosophical debate — i.e., aimed at the question of shared appearance between head-on ellipses and rotated circles. 5

To begin filling this gap, we ran a series of such studies (Morales et al., 2020, 2021). To acquire experimental evidence of shared appearance between head-on ellipses and

---

4Vision scientists also weigh in on this question; though their views may not perfectly map onto these philosophical positions, it is not difficult to find passages sympathetic to one or the other class. For example, Gregory (1970) writes: “[B]ehavior 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 object—table—not according to the brown patch in my eye”. Similar views are attributed to Attnave, who had a “sandbox in the head” theory of both perception and imagery: “These ‘sandbox in the head’ theories (Attnave, 1972) propose that depth is encoded directly in perceiving and imagining, and that no special status is afforded to the picture plane or any other two-dimensional surface” (Kerr, 1993). More recently, Erdogan and Jacobs (2017) present a model (also referenced by Burge and Burge) that seems aligned with this view in both the content and format of the representations involved: “Shape representations code information about an object’s three-dimensional structure, not the two-dimensional structure of its retinal image”, and “Shape representations code shape properties in an object-centered coordinate system, not a viewer-centered coordinate system”. This is not to say that such theories could not accommodate perspectival similarity or account for empirical evidence in its favor (nor that there aren’t other, contrasting scientific accounts), but rather that perspectival similarity is neither assumed nor actively predicted by these models. (Note that viewpoint-dependency in such models is not the same as perspectival similarity: viewpoint-dependency is the behavioral phenomenon whereby object recognition is better from some views than others, whereas perspectival similarity concerns shared appearance). For the sake of precision, and to avoid misrepresentation, the rest of our discussion focuses more on philosophical expressions of these views than scientific ones; however, we think many classical and contemporary scientific accounts express broadly similar theoretical positions.

5Many scholars have acknowledged this lack of relevant experimental work. For example, Schwenkler and Weksler (2019) note the traditional view that “the dispute belongs to the realm of phenomenological, conceptual, and metaphysical analysis”, rightly asserting that the debate has so far “proven recalcitrant to the methods of philosophy”. Reacting to this traditional view, they propose an experimental approach, and they (also rightly) consider the very idea of an experimental approach to be a new and different inroad to these issues (see also discussion in Green 2021; Kelly 2008). An intriguing exception might be Thouless (1931); in some ways our project is a continuation of this old but inspiring work.
rotated circles, we asked whether subjects who must locate a head-on ellipse in a display containing an additional non-elliptical object would be “distracted” by rotated circles more so than head-on circles, on account of their putatively shared appearance. This general approach and explanatory pattern is common in visual cognition research. For example, consider that it is harder to find a red square among red triangles than to find a red square among blue triangles (as reflected in slower search times). The canonical explanation of this pattern is that, even though red squares and red triangles look very different in some respects, they also share some aspect of their appearance (namely, their color). This shared aspect makes them harder to distinguish — especially under time pressure — than pairs of stimuli that do not share this or other aspects of their appearance (here, red squares and blue triangles). We reasoned that the same logic could hold here, if head-on ellipses and rotated circles look similar in a way that head-on ellipses and head-on circles do not.

In nine experiments reported in Morales et al. (2020), as well as a tenth reported in Morales et al. (2021), we indeed found this predicted slowdown: A head-on ellipse is harder to locate when displayed next to a rotated circular coin than a head-on circular coin. On the basis of this and similar results across many variations of presentation and context (including several variants in which the coins sat in front of subjects for seconds or even minutes), we concluded that “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.” Our follow-up publication refined this view further: “Our claim is that frontal ellipses compete with rotated circles because of some shared aspect of their appearance, where this aspect is shared from certain perspectives but not others” (Morales et al., 2021). In other words, we take these results to favor perspectival similarity, and thereby to oppose the view that there is not “any sense” in which a rotated circle looks similar to a head-on ellipse.

Recently, Burge and Burge (2022) issued a critique of this work, arguing that it “provides no empirical support for any claims regarding either perspective or the perception of shape” because of defects in our characterization of the philosophical and scientific views we discuss, flaws in our experiments (which, in their view, were “poorly controlled and poorly conceived”), and patterns in our data that allegedly undermine our conclusions. More specifically, Burge and Burge give three empirical arguments: (1) the reaction-time differences we focus on are also accompanied by accuracy differences, in ways that undermine our claims; (2) rotating a coin introduces new “uncontrolled differences” between it and the ellipse that could account for our results differently; and (3) the cues for discriminating distal circles from distal ellipses are weaker and less reliable when the distal circles are rotated.

This critique rests on a series of misunderstandings — of our experiments, claims, and, critically, the literature we are engaged with. Here, we respond to these objec-
tions on both philosophical and empirical fronts. In the first section (“The Evidence”), we directly address Burge and Burge’s empirical arguments and evaluate their success against our claims; not only do we find them unsuccessful, we show that many of them favor our account rather than oppose it (though we agree that the relevant issues are not fully decided or resolved). In the second section (“Perspectival similarity without ‘a new entity’”), we further clarify our position and address charges that we misrepresent the philosophical literature on this question. We show that, to the contrary, it is Burge and Burge’s philosophical discussion that is problematic, in ways that directly relate to their subsequent mischaracterization of our claims (which we detail and correct). This section also suggests that Burge and Burge’s positive view places them in much friendlier intellectual territory than their discussion suggests; it only seems otherwise because they ignore a prominent rival view that our own work was responding to (and that both our position and their position opposes). Finally, in the third section (“Moving Forward”), we point to future empirical directions on these philosophical issues, and reflect on ways to productively advance this literature.

1. The Evidence

Our research question is whether there is any shared appearance between a rotated circle and a head-on ellipse that is not shared between a head-on circle and a head-on ellipse. We test this question in 10 experiments that find impaired search for head-on ellipses flanked by rotated circles than by head-on circles (with some variations using squares and trapezoids). Burge and Burge raise three empirical objections against our experiments. Here, we proceed through these arguments and assess their quality.

Argument 1: Accuracy Differences and Alleged Dissociations

The primary measure we report in our paper is the difference in reaction time between the head-on-circle condition and the rotated-circle condition. Burge and Burge observe another pattern: In addition to an RT difference (faster with head-on-circular distractors than rotated-circular distractors), there is also an accuracy difference, such that subjects answer correctly more often in the head-on-circle condition than the rotated-circle condition. They refer to this pattern as “unreported data” and suggest that it undermines our arguments and conclusions.

However, this objection gets things backwards; the correlation between speed and accuracy complements our results, and in no way undermines them. As Burge and

---

"Readers are advised that every data point from every one of our experiments — along with our code, stimuli, and analyses — has been permanently and publicly available since publication (https://osf.io/thj6y/). Any reader is welcome, and indeed invited, to conduct any further analyses they wish."
Burge correctly note, speed and accuracy are joined at the hip; they are always correlated, except in very unusual circumstances — and so the question in any given case is only whether this correlation is positive or negative. The problematic pattern for our conclusion of shared appearance would have been a negative correlation between accuracy and speed: slower response times and greater accuracy in the rotated condition than the head-on condition. This pattern is known in psychophysics as a “speed-accuracy tradeoff”, and it makes RT differences difficult to interpret, because slower responses might not reflect increased difficulty or stimulus competition but rather strategic behavior by subjects to prioritize accuracy at the expense of speed.

But what Burge and Burge highlight is not a speed-accuracy tradeoff; it is the opposite pattern — a difference in accuracy that is in line with the RT slowdown. This is the pattern that visual cognition researchers hope to see in their data; it is a confirmation of the performance differences suggested by the RT results. That’s why visual search papers so often include phrases like “faster and more accurate” in their results sections; that phrase does not mean “faster and more accurate (and therefore uninterpretable)”; it means “faster and more accurate (and so especially compelling)”. (For examples of this canonical pattern of reasoning, see Kimchi 1994; Waters and Lipp 2008; Wolfe et al. 2011; there are hundreds of others.)

Indeed, Burge and Burge imply as much when they write that “there is nothing surprising” about this accuracy difference, whose relationship to speed “has been known for decades”. We agree. What is surprising is Burge and Burge’s suggestion that this undermines our arguments or conclusions. Recall: Our view is that there is some sense in which circles, when rotated, come to visually resemble head-on ellipses (relative to when the circles are viewed head-on). With this view in mind, note that an equivalent way of stating the accuracy difference is that subjects searching for ellipses under time pressure are more likely to respond to rotated circles than to head-on circles. This is exactly what one would expect on our view.

To see this another way, consider our earlier analogy to finding a red square among red triangles vs. a red square among blue triangles. Imagine we ran that toy experiment, such that on every trial, one location would have a red square and the other location would have either a red triangle or a blue triangle, with subjects simply having to locate the red square as quickly as possible. The expected results are straightforward: There would be an RT slowdown in the red triangle condition relative to the blue triangle condition, and there would also be an accuracy difference consistent with this effect. Its interpretation would be straightforward as well: The reason there would be slower responses and more mistakes for finding a red square beside a red triangle than beside a blue triangle is that red squares share an aspect of their appearance with red triangles that they don’t share with blue triangles. The same logic applies in our actual experiments. Perspectival similarity predicts the RT differences and the accuracy differences. But, recall, some prominent positions deny perspectival similarity, holding that there is not any sense in which a rotated circle
comes to share its appearance with an ellipse. So, while everyone agrees that red triangles are more distracting than blue triangles when looking for red squares, some views may not agree that rotated circles should be more distracting than head-on circles when looking for head-on ellipses. Our experiments disconfirm that hypothesis.

Why did Burge and Burge think this very ordinary result was a problem for us? As their text clarifies, it’s because they read us as claiming a dissociation between speed and accuracy, since we noted that our RT differences emerged even when “performance was at ceiling” (e.g., Experiments 8 and 9) and indeed even when subjects were “not confused at all about the true shapes of the stimuli”. But this is no claim of a dissociation between speed and accuracy. Of course, it is true that even in Experiments 8 and 9 performance was never literally 100%, and we are happy to accept Burge and Burge’s clarification that a subject who answers incorrectly on a given trial is in some weak sense “confused” about what they were looking at (though this same weak sense would apply equally to the red squares and triangles case, and so is no problem for our view). But once again these differences only further support, and do not in any way refute, our central claims. The only reason we mentioned high accuracy was to convey that the task was easy and the objects were trivial to distinguish (like red squares and red triangles). For example, in the strongest and most naturalistic experiment from our first publication (real-world objects with unconstrained viewing; Experiment 9), accuracy was 98.4% in the head-on condition and 97.2% in the rotated condition — extremely good (approaching “ceiling”) performance, especially in light of the time pressure. We mentioned this only because, if accuracy had been poor (say, 70%), then perhaps an argument could be made that we had ineffectively displayed our stimuli. But 98.4% and 97.2% accuracy, under time pressure, is remarkably good, suggesting that the objects were easily seen for what they were: circles and ellipses presented at various orientations. We did not claim a dissociation between speed and accuracy. We were only arguing that the reaction time differences arose even under conditions where accuracy was high. All of these results straightforwardly support our account, and in no way undermine it.

**Argument 2a: New Uncontrolled Differences**

Burge and Burge’s other arguments concern what they call “uncontrolled differences” in our two experimental conditions (head-on circular distractor and rotated circular distractor). There are two variants of this argument in their text, so we address them separately here.

---

7Burge and Burge describe the data from our real-world experiments as “filtered”. This was true of our computer-based experiments (where we applied an 80% accuracy exclusion criterion to ensure data quality, though all results remained the same even without this filtering); but it was not true of our real-world experiments, where every subject in all three of these experiments performed well enough that not a single subject approached our accuracy cutoff. Additionally, all subjects passed a screening procedure that required 100% shape identification performance without time pressure.
First, Burge and Burge note that, when a coin is rotated, other visual differences come along with this manipulation, such as increased visibility of the coin’s rim, a change in its specularity and shading, and so on. They suggest that, rather than reflecting perspectival similarity, results in the rotated-circle condition (i.e., worse / slower performance) could have arisen because “subjects could identify the distal circle as the stimulus with the darker shading, or with the visible edge, among other cues”. On this worry, the more such differences there are, the less interpretable our experiments are. For example, our real-world experiments further amplify the problem, because “these experiments introduced more uncontrolled cues (binocular disparity & motion parallax)”. But this objection too gets things backwards; it is the opposite of the argument Burge and Burge need to make to undermine our claims. It is true, of course, that rotating the circle introduces many new differences between the circle and the ellipse. (Most of these differences would more ordinarily be called cues to 3D shape.) But these new differences — differences between the coins that are added in the rotated-circle condition — should make our result more compelling, not less compelling. To reiterate: Our finding is that locating the head-on ellipse is more difficult in the rotated-circle condition than the head-on-circle condition — it produces slower responses (and less accurate ones, as Burge and Burge note). Our explanation of these performance differences is that a rotated circle and a head-on ellipse look similar in a way that a head-on circle and a head-on ellipse do not, and that this makes the rotated-circle condition harder than the head-on-circle condition. If Burge and Burge’s objection is to succeed, it needs to give an alternative account of why that condition is difficult. Pointing to new or additional differences between the stimuli in the rotated-circle condition seemingly does the opposite of what Burge and Burge need to do; each additional difference on their list is a difference that would tend to aid performance in that condition, not impair it. The newly visible rim (on one coin but not the other), the new differences in shading and specularity between the coins, and so on, are all features that make targets and distractors more discriminable from one another, not less (since, e.g., anytime a subject sees a rim, they could reject that shape as their target).

To see this another way, consider Burge and Burge’s observation that removing all low-level differences between the coins in the rotated-circle condition would have made the task impossible. That is surely correct. But this reasoning extends in the other direction too: Just as removing differences between the coins in the rotated-circle condition would have made the discrimination even harder (and eventually impossible, if all differences were removed), the presence of those differences should tend to assist the discrimination (barring some independent reason to think that shading/specularities/disparities/etc. of one kind but not the other impair discriminations). And so the presence of those additional differences between the coins does not immediately provide an alternative explanation for why the rotated-circle condition is the harder
condition. It is possible that, for the claims Burge and Burge imagine we are making (see next section), this objection would be more powerful or effective. But for our claims of perspectival similarity, their objection is pointing in the wrong direction.

**Argument 2b: Unreliable Cues to Distal Shape**

Burge and Burge’s final argument may be their strongest one (in part because it is the only one whose upshot would be a problem for our view). This argument is also based on differences in the head-on vs. rotated condition. Here, Burge and Burge argue that “Distal circles and ellipses that generate retinally projected shapes with matched aspect ratios force subjects to rely more on other cues (e.g. shading, specularity, edge visibility)”, and that these other cues “are as a group weaker and less reliable”. Since “Less informative, less reliable cues tend to cause slower response times and less accurate discriminations”, it is possible that the slowdown we observed occurred only because it is harder to recover the distal shape of the rotated circle (in ways that allow for discrimination from distal ellipses) than the distal shape of a head-on circle. This concern is closely related to Linton (2021)’s worry, namely that distal shape was more difficult to recover in the rotated condition.

If that difficulty explains our results, this would indeed be a problem for our view. What do we make of this argument? First, we note that Burge and Burge’s rendition of this objection is difficult to evaluate as written, in part because they offer no sourcing for their claims — i.e., their claim that the cues available in the rotated case are, as a group, weaker and less reliable than in the head-on case, and their claim that these cues tend to cause slower response times. (Indeed, there are no citations, in either of the paragraphs where this argument is given, to any sources at all.\(^8\))

Nevertheless, we can consider it. And indeed we did. This general concern motivated nearly half of our experiments, which address it from multiple angles:

- Experiment 6 introduced a delay in responding, such that subjects had to view the stimuli for a full second before being permitted to give a response (with

\(^8\)The objection is also given by Burge and Burge in a way that seems to concede an important aspect of our claim. Why do “distal circles and ellipses that generate retinally projected shapes with matched aspect ratios force subjects to rely more on other cues”\(^?\)? Similarities or differences in retinally projected shape are not determinative of similarities or differences in distal shape: As Burge and Burge (and all students of visual perception) well know, underspecification of the distal environment by the retinal image entails that two objects with different projections can have the same distal shapes, and two objects with the same projections can have different distal shapes. This means that subjects in both the head-on-circle and rotated-circle conditions were “forced” to rely on other cues to distal shape, because different retinal projections are consistent with similar distal shapes (just as similar projections are consistent with different distal shapes). In other words, while we, the experimenters, know that the object with a circular projection was in fact a head-on circular object, the naive subject’s visual system must discover this fact independently: the circular projection of the head-on circle is perfectly consistent with a cleverly angled elliptical object.
those responses thus arriving over 1.5 seconds after exposure to the stimuli). So even if there are differences in the strength of the cues to distal shape in each condition (whatever the direction of these differences might be), subjects in this experiment should have had enough time to adequately recover distal shape. But if perspectival similarity persists in experience even after shape constancy is achieved, we might still predict an impact on search behavior. That in fact is what we found.

- Experiments 8 and 9 used real-world objects in full-cue conditions, and the objects remained in front of subjects for nearly half an hour. The coins were visible throughout the entire session (including before the first trial), under more than adequate illumination, with no masking. In other words, the setup was just an array of easily visible wooden objects, located right in front of the subjects. Unsurprisingly, 100% of subjects in those experiments passed a screening procedure requiring them to say which shapes were which, perfectly, without time pressure. Still, our effects remained.

- Finally, in response to Linton (2021)’s related objection, Experiment S1 (reported in our follow-up publication; Morales et al. 2021) replicated the setup of Experiments 8 and 9, but at a viewing distance that has been shown in other work to be optimal for shape constancy from binocular disparity (Johnston, 1991). It is hard to imagine better and more ecologically valid conditions for perceiving the distal shape of an object; it seems to us that the only (or at least most plausible) reason for a slowdown under these conditions is if the two objects bear a representational similarity to one another.

Burge and Burge briefly discuss these experiments, but only to dismiss them on the basis of the other, flawed arguments reviewed earlier. For example, of Experiment 6 (delay), Burge and Burge write “these results also exhibit accuracy differences”; and of Experiments 8, 9, and S1 (real-world), they write “But these experiments introduced more uncontrolled cues (binocular disparity & motion parallax). And, here too, there are accuracy differences between hard and easy conditions”. As we have seen, these objections are not, and cannot be, decisive. Accuracy differences are expected whenever there are RT differences, and they are most definitely predicted on our account; appealing to them again here does not undermine these experiments as answers to the “unreliable cues” objection. And Burge and Burge’s concern about uncontrolled cues again goes in the wrong direction: The addition of differences in the rotated-circle conditions (“these experiments introduced more uncontrolled cues”) is no objection at all when the rotated-circle conditions are the harder conditions. Of course, we do not mean to say that our results fully decide this issue; we think they are highly suggestive when considered together (and in conjunction with our six other
experiments, which included an experiment where height and width were varied independently of shape [Experiment 2], an experiment that further explored the role of rotation, specularity, and edge visibility [Experiment 3], and an experiment that matched projected aspect-ratios, but not projected shapes, in all conditions [Experiment 4]), and we are very much open to other approaches and sources of evidence (see final section). But Burge and Burge need new responses to these experiments.

2. Perspectival similarity without “a new entity”:

Do we (dis)agree about perspectival shape?

Its tone aside, one of the more surprising aspects of Burge and Burge’s article is that it eventually expresses a view that sounds quite congenial to our own. In a passage that arrived unexpectedly following their vigorous discussion, Burge and Burge write: “we are commonly aware of some elliptical shape corresponding to the projection cast by a rotated dinner plate”. This statement sounds very much like our own claims, and Burge and Burge’s efforts to distinguish the two are subtle indeed:

[P]rojected shape is merely sensed, but not perceived. The elliptical-shape-awareness depends causally on this mere sensing. It is an open empirical question at what processing stage—mere sensory, perceptual, or post-perceptual—this conscious awareness emerges and what psychological capacities it requires. We think it likely that conscious mere sensing of retinally projected shape occurs at the perceptual stage of processing.

So: For Burge and Burge, the elliptical aspect of the rotated coin is sensed but not perceived (not in itself an unusual view on this question), yet the awareness that derives from this non-perceptual sensing likely arises at the “perceptual stage of processing” (as opposed to the “mere sensory” stage), and is consciously accessible. That is more than sufficient for our purposes as laid out above. Whether the coin’s ellipticity is visually apparent, visually perceived, seen, or consciously-sensed-at-the-perceptual-stage-of-processing, all such formulations disfavor views that reject perspectival similarity and favor views that embrace perspectival similarity — and so are well in line with the overall aim of our project.

What, then, explains our apparent disagreement? We think the answer lies in another surprising aspect of Burge and Burge’s discussion: It does not actually engage with (or even mention) the philosophical views that our article was responding to — namely, the class of views that reject perspectival similarity. Note, for example, that the philosophical sources we rely on most heavily in our original article — the work by Schwitzgebel and Smith reviewed above (which are the only pieces of contemporary philosophy quoted in extenso in our text, and which appear very prominently in our
introduction, immediately before we introduce our empirical approach) — are not
discussed at all in Burge and Burge’s own piece. They do not attempt to show that
we have misunderstood these views, nor those of any other philosopher whose work
we discuss. Indeed, their article does not once cite these pieces of work that were so
central to our project.  

We think this omission matters, because it leads Burge and Burge to misinterpret
our claims and criticize views we do not hold. Instead of reading us as other scholars
have (e.g., Cheng, 2022; Cohen, 2021; Daoust, 2021; and especially Green 2021, who
reviews the same sources and context and then characterizes our views accurately
and as intended: “visual experiences of the slanted circle and head-on ellipse are
similar in some respect”), Burge and Burge exoticize our claims, attributing to us the
proposal that “a new entity” should be introduced into the science of perception and
interpreting this entity as being located “in some intermediate position(s) between
the distal shape and the retinal image”. This is not our claim. It is true that some
philosophers sometimes conceive of perspectival appearance this way (perhaps, e.g.,
in discussions of “Thouless properties”; Hill, 2016), but we are not among them.
Our text does not defend this conception. Indeed, in an effort to be cautious and
conservative about this general issue, we wrote:

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 is-

quisitions at stake here, it may be important to distinguish between interference
caused by matching perspectival shapes vs. by persisting retinal images

---

9 For their part, Burge and Burge bring a similar charge against our discussion, arguing that
“Morales et al. barely discuss philosophy that utilizes science’s treatment of perspective in percep-
tion”, offering instead the following five publications: Burge (2010); Burge (2014a); Burge (2014b);
Lande (2018); and Rescorla (2014). In fact, our article discussed or referenced nearly two dozen
works of philosophy, among them the following publications that are substantively engaged with the
science of perception (in general) and visual perspective (in particular): Bennett (2012); Bennett
(2016); Briscoe (2008); Cohen (2010); Green and Schellenberg (2018); Hill and Bennett (2008); Noé
(2004); Weksler (2016); and Wojtach (2009). Burge and Burge’s charge may be related to their
surprising dismissal of work by Cohen (2010) and Hill (2016), asserting or implying that it is “not
scientifically informed”. Cohen’s paper in particular is remarkably well-integrated with relevant sci-
entific literature (and more generally, Cohen and Hill are model examples of scientifically informed
philosophy of perception). Moreover, Burge and Burge’s list of alternative references is misleading
on this score: Most items on it (Burge, 2014a; Burge, 2014b; Rescorla, 2014) are contributions to
author-meets-critics discussions that discuss no relevant empirical work on visual perspective as it
relates to shape perception (though some are empirically engaged in other ways); indeed, Burge
(2014b) contains no references to any scientific work at all. Burge (2010) is the second author’s
monograph exploring objectivity in perception and cognition; we agree that it is a relevant source,
and we regret having failed to cite it. Lande (2018) is by far the most relevant work on Burge and
Burge’s list, and it is accordingly cited multiple times in our paper.
themselves vs. by independent representations of ellipticity [Lande, 2018].
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.

In other words, we are explicitly uncommitted to the view Burge and Burge attribute to us. 10 Of course, it is possible we could have been even clearer about this, and we acknowledge Burge and Burge’s concern that we sometimes express “ambiguous, unclear theoretical positions”. (For example: “Throughout their article, Morales et al. shift between using the term ‘perception’ and using the terms ‘perceptual experience’ or ‘conscious experience’.”) Though other scholars interpreted us correctly despite this ambiguity (and we are aware of no scholars who read us as Burge and Burge did), we nevertheless embrace this feedback and aspire to articulate ourselves more clearly in the future. But it is also important to consider claims with respect to the context in which they appear. We hope we have now clarified that our work is engaged with the (unresolved) debate over whether there is perspectival similarity in visual appearance. 11

Indeed, on an accurate reading of our claims, Burge and Burge not only express sympathy towards them but actively endorse them. They write: “If the authors were to claim merely that there is a viewpoint-dependent shape-similarity in how distal shapes are perceptually represented (or in how they appear), the authors would be stating something that is already well-appreciated”. But this is our claim. (Again: “Our claim is that frontal ellipses compete with rotated circles because of some shared aspect of their appearance, where this aspect is shared from certain perspectives but not others.”) Burge and Burge are simply incorrect, as a matter of scholarship, that

---

10 Burge and Burge quote this passage too, but use its cautiousness against us. They write that we “clearly distinguish perspectival shape from the retinal image” and so they take license to attribute to us a much stronger view in which perspectival shape is separated from any retinal or projective properties. But in fact we distinguish these views precisely to clarify what we are not claiming, since we are aware that some philosophers do argue for perspectival properties conceived along these lines. We do not conclude that our results support an intermediate entity located between the perceiver’s eyes and the world; indeed, we go out of our way not to draw this conclusion.

11 Burge and Burge’s mishandling of our views complicates other aspects of their discussion. For example, they report that they contacted multiple vision scientists whose work we quote in our paper and found that “In each case, they hold views very different from the present authors’ portrayals”. This was already a somewhat puzzling statement given that only one of the works of vision science quoted in our main text has any living authors. (Indeed, the maximum number of quotations that Burge and Burge could be referring to is two: the passage from Murray et al.’s neuroimaging study, where Burge and Burge agree that our reading is “perhaps invited by some of Murray et al.’s terminology”, and a passage from a textbook appearing in a footnote.) And Burge and Burge provide no details about the questions they asked or the answers they received. In any case, given their misconstrual of our aims, it is not clear what to make of their reporting. (Indeed, we regret that we were not among the scholars Burge and Burge contacted for this sort of clarification.)
perspectival similarity in perception (or appearance) is “already well-appreciated”, if by well-appreciated they mean uncontroversial, widely accepted, or any meaning in that vicinity. As we have shown in our original publications and again here in this discussion, many prominent contemporary scholars hold views that reject this supposedly well-appreciated claim, in clear writing that has been widely interpreted in the literature along these lines. These are the same views and sources that Burge and Burge omit from their critique. If Burge and Burge find our claim obvious, so be it; this means only that our experiments were not designed for them. If anything, they or anyone else with that feeling could see our results as complementary to their views — i.e., as providing additional empirical support for a position they accept but that others reject (just as our contribution has been received by other scholars already sympathetic to perspectival similarity; e.g., Green, 2021).

3. Moving Forward

Our experiments are in no way the last word on this deep, longstanding, and foundational issue about the nature of visual perspective. Indeed, one could even say they are among the first words in a new empirical conversation about this question (one perhaps started by Kelly 2008, and Schwenkler and Weksler 2019). How might that conversation productively move forward, both philosophically and empirically?

One issue that remains unclear (both in Burge and Burge’s discussion, and in the literature more generally) is the extent to which the present questions are ultimately

---

12Our interpretation of Schwitzgebel’s and Smith’s views and their place in the literature is mainstream, shared by multiple scholars writing before and after us (e.g., Cohen, 2021; Green, 2021; Schwenkler and Weksler, 2019). For example, Green (2021), reviewing this work, describes perspectival similarity as an active, “controversial” question in the literature and characterizes these authors’ views as follows: “[T]he slanted circle does not appear similar in any relevant respect to the head-on ellipse (Smith, 2002; Schwitzgebel, 2006, 2011; Hopp, 2013). […] This account does not predict any similarity between experiences of the slanted circle and the head-on ellipse.” Similarly, Cohen (2021) describes the work of these same three authors as arguing that we see “roundness exclusively” when we look at rotated circular coins. We read these sources as Cohen and Green do.

13Despite this agreement on the question of whether there is viewpoint-dependent (i.e., perspectival) similarity in these cases, we understand that there may be differences in how we and Burge and Burge understand the nature or basis of this similarity. For example, their use of the terms “perception” and “perceptual representation” seem to be heavily informed by the framework developed, most prominently, in Burge (2010) — i.e., as a process of “objectification” marked by perceptual constancies. Suggested in their response too is a distinction between what is represented in perception and how it is represented, which has philosophical origins in the Fregean notion of “modes of presentation”. However, we do not further pursue these distinctions here because (a) they are secondary to our own question of whether there is perspectival similarity in the representation of distal shape (where both we and Burge and Burge answer yes, while other scholars answer no), and because (b) Burge and Burge do not further explicate these distinctions in their response, and we do not want to risk misinterpreting them.
empirical (or can be informed by empirical data). For some issues, Burge and Burge seem to suggest not. For example, when discussing our paradigm, they not only note uncontrolled low-level image properties (though, as we argue above, these are characteristics that favor our interpretation rather than oppose it) but also insist that the stimuli were “uncontrollable” — i.e., that no other version of this approach could have succeeded in generating relevant data for these questions. It is far from obvious that this follows. Burge and Burge seem to be asking for a single “critical experiment” without any possible confounding factors. But many substantive views about perception (or any other topic, for that matter) are supported not by a single critical result but rather by evaluating how they cohere with an overall body of evidence, most individual pieces of which will by necessity be incomplete and imperfect — but not, just for this reason, irrelevant. If, in their view, no experiment like ours could provide “any evidence” bearing on these questions, then what kind of results would? (Burge and Burge do propose one experiment in a certain species of fish; but this seems to rest on some of their earlier misunderstanding of our claims.) In any case, we do not share Burge and Burge’s pessimism about the scientific tractability of these questions. At the very least, we believe our more modest claims about perspectival similarity are amenable to empirical investigation — including, but not limited to, our own methodological approach.

We are encouraged to find our own empirical optimism echoed elsewhere in this literature. For example, Kelly (2008) reports pilot data from a priming experiment asking whether viewing rotated circular objects facilitates subsequent recognition of drawn ellipses. Complementarily, Schwenkler and Weksler (2019) propose an experiment to adjudicate perceptual vs. post-perceptual interpretations of perspectival representation (rather than sensory vs. perceptual interpretations), by varying working memory demands in tasks that ask subjects to report perspectival shape properties. Linton (2021), even while criticizing our work, shares a creative design using an elliptical cloud of points whose shape is visually distorted in different ways under monocular and binocular viewing conditions. Finally, Cheng (2022) outlines a neuroimaging approach based on repetition suppression. Still other empirical work could more directly explore the role of conscious awareness in these phenomena and results, which we discuss only briefly in our own work but which merits independent investigation (see, e.g., discussion in Cohen, 2021; Green, 2021; Morales, 2021).

Each of these latter proposals seems promising to us. Just as importantly, they seem productive — and, especially, constructive. Regardless of whether these authors broadly share our views, express some skepticism, or are genuinely undecided, they have embraced the sorts of empirical approaches we have promoted ourselves, and they point forward in ways that will surely enrich a philosophical debate that has at times risked becoming stale and intractable. We hope the present discussion finds a similarly constructive course.
References


