

# Seeing the Tipping Point: Balance Perception and Visual Shape

Chaz Firestone and Frank C. Keil  
Yale University

In a brief glance at an object or shape, we can appreciate a rich suite of its functional properties, including the organization of the object's parts, its optimal contact points for grasping, and its center of mass, or balancing point. However, in the real world and the laboratory, balance perception shows systematic biases whereby observers may misjudge a shape's center of mass by a severe margin. Are such biases simply quirks of physical reasoning? Or might they instead reflect more fundamental principles of object representation? Here we demonstrate systematically biased center-of-mass estimation for two-dimensional (2D) shapes (Study 1) and advance a surprising explanation of such biases. We suggest that the mind implicitly represents ordinary 2D shapes as rich, volumetric, three-dimensional (3D) objects, and that these "inflated" shape representations intrude on and bias perception of the 2D shape's geometric properties. Such "inflation" is a computer-graphics technique for segmenting shapes into parts, and we show that a model derived from this technique best accounts for the biases in center-of-mass estimation in Study 1. Further supporting this account, we show that reducing the need for inflated shape representations diminishes such biases: Center-of-mass estimation improved when cues to shapehood were attenuated (Study 2) and when shapes' depths were explicitly depicted using real-life objects laser-cut from wood (Study 3). We suggest that the technique of shape inflation is actually implemented in the mind; thus, biases in our impressions of balance reflect a more general functional characteristic of object perception.

*Keywords:* visual cognition, balance perception, shape perception, center of mass, inflation

*Supplemental materials:* <http://dx.doi.org/10.1037/xge0000151.supp>

Whereas many striking visual phenomena—including nearly all classical visual illusions—require impoverished viewing conditions, tricks in lighting, carefully crafted stimuli, or accidental views of precisely arranged scenes, a more familiar and naturalistic sort of illusion involves the (mis)perception of physical balance. In certain cases, we may observe one or more objects appearing to balance in circumstances that seem unlikely or even impossible, as when a ballet dancer holds a seemingly off-tilt pose, a stage performer stacks several objects on her head, or an architectural design gives a building the false appearance of being about to topple.

Thus, perfectly balanced objects can falsely appear unbalanced. In addition, in many cases, we may illusorily perceive not only that an object is unbalanced but also that it seems unbalanced in a particular way and might therefore appear as though it should fall over in a particular *direction*. For example, a popular children's toy sometimes known as a "gravity bird" can balance perfectly on the tip of its beak in apparent defiance of physics, seeming as

though it should fall backward toward its tail (Figure 1a). Likewise, pieces of rock-balancing art can involve surprisingly stable arrangements of stacked heavy objects and yet give the illusion that the rocks should topple over in a particular direction. For instance, Figure 1b shows an image of a stable tower of rocks, but the tower seems unbalanced and likely to fall (perhaps to the right).

## (Mis)perceiving Balance: A Puzzle

How and why does our perception of balance go wrong? Here we explore a curious bias we discovered in the laboratory that mirrors the sorts of directional biases you may experience in Figure 1. To preview the initial result, subjects who estimate the center of mass of a wide variety of two-dimensional (2D) shapes (by identifying the point on the surface of a shape at which the shape would be able to balance) give surprisingly inaccurate estimates, often misjudging the center of mass by a wide margin. Moreover, subjects' errors for a given shape are not randomly or normally distributed; instead, their error from the true center of mass of a particular shape is consistently biased in a particular direction, such that even the location converged upon by the average of many subjects' estimates is itself often far off-target (see Figure 2).

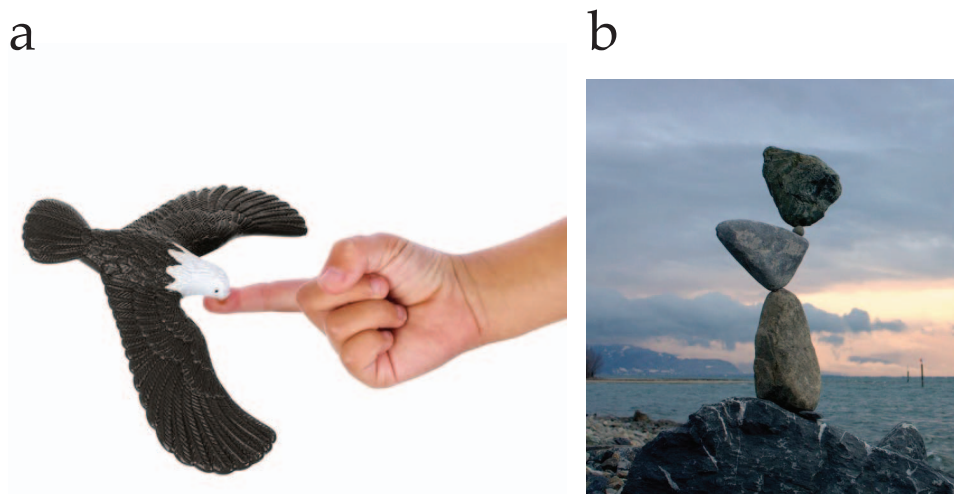
Such systematic biases in center-of-mass estimation (which are empirically demonstrated and characterized in detail in Study 1) would seem surprising for several reasons. The appreciation of such physical properties of objects seems crucial to many familiar activities that we apparently perform successfully (e.g., hanging picture frames, carefully stacking dishes in a sink, or adding blocks to towers as children), including actions that would have been

---

Chaz Firestone and Frank C. Keil, Department of Psychology, Yale University.

For helpful conversation, the authors thank Jacob Feldman, Melissa Kibbe, Brian Scholl, Manish Singh, and members of the Yale Perception and Cognition Laboratory and the Yale Cognition and Development Laboratory.

Correspondence concerning this article should be addressed to Chaz Firestone, Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205. E-mail: [chaz.firestone@yale.edu](mailto:chaz.firestone@yale.edu)



*Figure 1.* Examples of real-world illusions of balance. (a) A toy sometimes known as a “gravity bird” balances on its beak in apparent defiance of physics. (b) Stable arrangements of rocks can falsely look as though they should topple over. See the online article for the color version of this figure.

important in our evolutionary past (e.g., lifting heavy objects, building simple structures, or rapidly grabbing and using tools; Iberall, Bingham, & Arbib, 1986) — and so it would be odd for our judgments of such properties to be so systematically inaccurate. In addition, the center of mass for shapes such as those in Figure 2 is simply the centroid of each of the shape’s pixels, and centroid estimates *are* accurate in other contexts (e.g., in perceiving the average location of many dots; Alvarez & Oliva, 2008), which suggests that the visual system can in principle accurately compute the center of mass of a 2D shape but for some reason does not do so. Finally, the existence of a directional bias in center-of-mass estimates, rather than normally distributed error around the true center of mass, is unusual: Even if the task is difficult for each subject, one might expect center-of-mass estimation to exhibit a “wisdom of the crowds” pattern in which random subject-level error would converge on the correct response (Galton, 1907). However, in this case, not only does each subject give inaccurate center-of-mass judgments, but the crowd is also “unwise.” Thus, this bias poses a puzzle: What is the nature of our poor perception of balance, and why does this bias arise?

### Shape “Inflation”: A Solution

Here we advance a surprising solution to the puzzle of directional biases in center-of-mass estimation. We suggest that these biases are not simply quirks of physical reasoning but that they instead reflect more general principles of object representation and perceptual organization. In particular, we suggest that such biases arise in the course of the visual segmentation of shapes into parts; for example, as when we perceive that a human hand is not a single unitary object but is rather composed of five distinct fingers and a palm (De Winter & Wagemans, 2006; Hoffman & Singh, 1997). Identifying the parts of a 2D shape is a fundamental and yet notoriously difficult computational process—indeed, much more difficult than identifying the parts of a three-dimensional (3D) shape (Hoffman & Richards, 1984). For this reason, a prominent approach to 2D part segmentation in computer vision models is to

“inflate” 2D shapes into a 3D form so as to apply more traditionally successful 3D segmentation rules (e.g., the 3D Minima rule; Hoffman & Richards, 1984) and then to project the part boundaries output by these 3D-specific rules back onto the 2D shape (Twarog, Tappen, & Adelson, 2012).

Thus, an intriguing possibility is that human vision too implicitly represents 2D figures in an “inflated” 3D form—even when no depth information is explicitly present in the image—and that this 3D representation plays a role in 2D shape representation, intruding on perception of the shape’s 2D properties (such as its center of mass) and biasing them toward that property’s value for the 3D object. Such an account would suggest that biased estimation of the 2D center of mass actually reflects a more general functional characteristic of object processing, and it would also constitute the first psychophysical evidence that shape inflation is not merely an effective technique in computer vision systems but is actually implemented in the mind.

### The Current Studies

In the present experiments, we first establish these systematic biases in center-of-mass estimation and compare the “inflation” hypothesis against other possible accounts (Study 1). Then, we directly test the inflation hypothesis by “disrupting” the hypothesized inflation process in two highly divergent ways. First, we degrade the shapes by presenting them as 2D arrays of dots, predicting that such degradation will (rather counterintuitively) *improve* center-of-mass estimation by reducing the impetus for such inflation in the first place (Study 2). Then, we enhance the shapes by presenting them as real-life, uniformly thick 3D objects (custom-constructed out of wood using a laser cutter), predicting that this will also improve center-of-mass estimation by reducing the need to rely on an inferred 3D representation of the 2D shape because such 3D information is already provided explicitly by the object itself (Study 3). Collectively, these studies test the possibility that the visual system builds rich, volumetric 3D models of

simple 2D shapes and that these inflated 3D representations drive basic judgments about objects in the world.

### Study 1: Inflationary Intrusion on Center-of-Mass Estimation

To establish and quantify systematic biases in center-of-mass estimation, we displayed asymmetric 2D shapes to subjects and instructed the subjects simply to indicate where on each shape's surface it could stably balance. We then compared these biases to a simple inflation-based model and directly tested it against other interpretations of such biases.

#### Method

**Subjects.** One hundred subjects were recruited online from Amazon Mechanical-Turk and were monetarily reimbursed<sup>1</sup> (for discussion of this subject pool's reliability, see Crump, McDonnell, & Gureckis, 2013).

**Stimuli.** Thirty shapes served as experimental stimuli (see Figure 2). Each shape met the following criteria: (a) the shape was asymmetrical about a major axis; (b) the true center of mass was located within the shape's bounding contour; and (c) the shape had a salient "part structure," such that certain regions of the shape were relatively long and thin whereas others were relatively short and stout. The shapes ranged from 362 to 637 pixels in width and 255 to 655 pixels in height (although their physical size was of course determined by the settings of each subject's display). All shapes appeared as black outlines with gray fills on a white background. (To ensure that subjects understood the task, we also included three "catch" shapes whose centers of mass we predicted would be easily and accurately estimated: an equilateral triangle, an ellipse, and a symmetrical "dumbbell" shape.)

**Procedure.** An instruction page informed subjects that they would see several shapes on subsequent pages and that they should "find (and then click) the 'center of gravity' of each shape. In other words, you should click the point on the surface of the shape where someone could balance the shape on their finger." Subjects were then shown an example of a correct response for a simple case (a square) before being reminded again that the task was to click "the point on the figure where it would be perfectly balanced." Each shape was then displayed one at a time on its own page in a randomized order between subjects. There was no time limit on responses.

#### Results and Discussion

Center-of-mass judgments for the experimental shapes were systematically biased (see Figure 2). Individual subjects were prone to error not only individually (deviating from the true center of mass by an average of 60 pixels, or 16% of the maximum error that would be possible even in principle given the constraints of the task) but also as a group: As plotted in Figure 2, the average location of subjects' center-of-mass estimates for a given shape itself deviated from that shape's true center of mass (i.e., the mean of the coordinates of the shape's interior points<sup>2</sup>) by an average of 43 pixels across shapes, reaching as high as 88 pixels (or 26% of the maximum possible error) for one shape. In other words, the "crowd" itself was unwise (cf. Galton, 1907), such that subjects'

estimates were individually and collectively often very far from the true center of mass. (By contrast, performance on the catch shapes was accurate: Across the three shapes, the average location of subjects' estimates deviated by only 7 pixels from the true center of mass, which suggested that subjects understood the task as intended.)

**Inflationary intrusion.** Closer inspection of the plotted center-of-mass estimates revealed a seemingly consistent pattern in subjects' errors: The estimates seem to "underweight" the contribution of long, narrow shape-parts to the overall center of mass and correspondingly "overweight" the contribution of wide, stout shape-parts. For example, estimates for the first shape depicted in Figure 2 (top row, leftmost column) illustrate this pattern: The average estimate (represented by the blue circle) is farther from the narrower parts (and closer to the wider parts) than is the true center of mass (represented by the green star). Such biased weighting may not be entirely irrational: Indeed, for many real-world objects—such as trees, pears, or bottles—the parts of the object that appear wider in 2D cross-section (e.g., the tree's trunk) are also in fact thicker in three dimensions relative to the parts of the object that appear narrower in two dimensions (e.g., the tree's branches), which are often thinner in three dimensions. Perhaps, then, subjects' estimates reflected an implicit assumption that wider shape-parts are actually thicker (and therefore heavier) and that narrower shape-parts are actually thinner (and therefore lighter), although such depth information was not explicitly present in the image (cf. Cole et al., 2009; Koenderink, van Doorn, & Kappers 1992).<sup>3</sup>

Conveniently, "inflationary" approaches to shape processing in the computer vision literature have recently formalized exactly this assumption. We naturally perceive shapes as having parts (Hoffman & Singh, 1997); thus, a longstanding challenge for theories of image processing has been to characterize how the visual system determines this part structure (including, e.g., how many parts there are, where one part ends and another begins, and so on). Traditional techniques for such 2D shape segmentation typically establish part boundaries along minima of concave curvature, but such approaches are frequently misled by common cases, such as shapes with undulating bounding contours or even a human hand (see Twarog et al., 2012). In more recent inflationary approaches, 2D images of shapes are segmented by first deriving 3D versions of them and then applying more successful 3D segmentation rules before projecting those part boundaries back onto the 2D shape. For example, one prominent and elegant approach known as *puff-*

<sup>1</sup> Given the exploratory nature of this study and the ease of online data collection, we chose this sample size simply because it seemed very large for a visual perception study of this sort. (For example, it was 10 times as large as that of the one previous study we know of using a similar task; Baud-Bovy & Soechting, 2001).

<sup>2</sup> Although 2D shapes do not literally have mass (and so, in a sense, had no "true center of mass" in this study), it is standard in this literature to refer to 2D shapes in this way (e.g., Baud-Bovy & Soechting, 2001; Cohen & Singh, 2006; Samuel & Kerzel, 2011). Throughout this paper, any instance of "true center of mass" could be read as "mean of the coordinates of the shape's interior points."

<sup>3</sup> Although subjects did not explicitly report making this nonuniform-thickness assumption in debriefing when asked to describe their estimation strategy. (If anything, many subjects reported assuming uniform thickness, describing their strategy as imagining "a piece of cardboard" or "a piece of paper cut into that shape.")

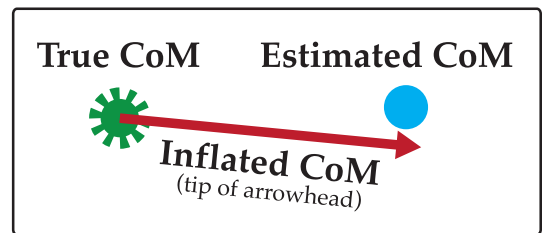
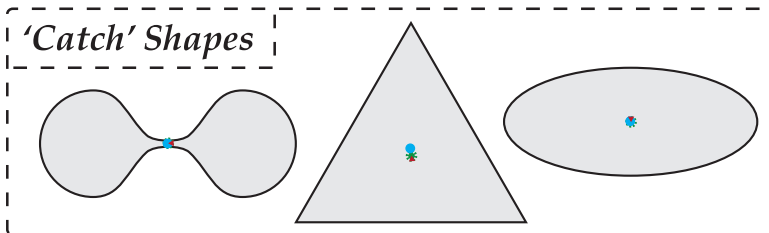
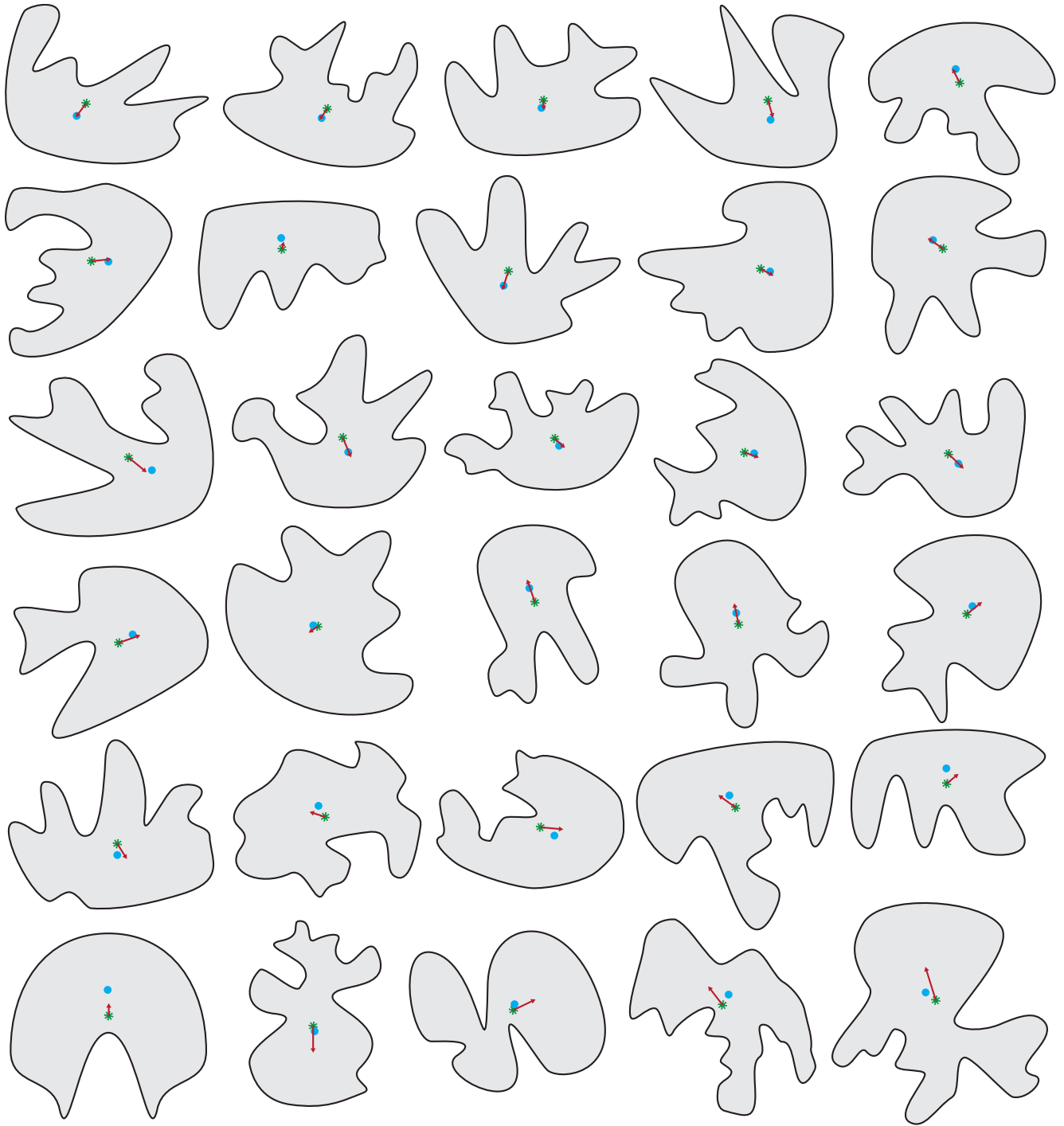


Figure 2 (opposite).

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

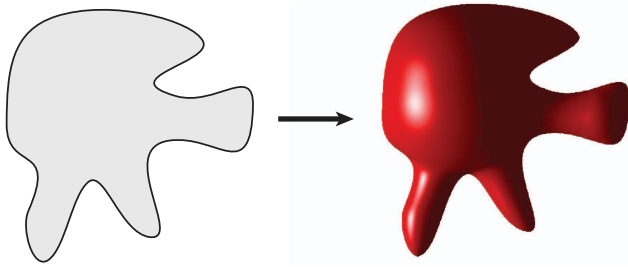


Figure 3. Example of Puffball inflation for a 2D shape used in the present studies. See the online article for the color version of this figure.

ball (Twarog et al., 2012) inflates 2D shapes by placing spheres of radius  $r$  anywhere inside of the 2D shape that a circle of radius  $r$  can be placed, and then taking the union of the spheres. (More pithily: “wherever you can place a circle, place a sphere.”) The resulting balloon-like objects (for an example, see Figure 3) can then be segmented into parts more easily (e.g., by using points of minimal principal curvature).

Crucially for explaining the presently discovered biases, (a) such inflation incidentally implements the principle that wider shape-parts are thicker and narrower shape-parts are thinner because the thickest regions of the inflated 3D shape (i.e., the points at the centers of the largest spheres) are necessarily the widest regions of the original 2D shape (i.e., the points that are the centers of the largest circles) and (b) this inflation process is based on principles of visual processing that are already known to be implemented by the mind (rather than appearing only in empirically unmotivated computer vision models), such as the deployment of medial-axis representations of shapes (Firestone & Scholl, 2014; Kovács, Fehér, & Julesz, 1998; Kovacs & Julesz, 1994; Wilder, Feldman, & Singh, 2011).

Thus, as an initial test of whether biases in center-of-mass estimation could be explained by intrusion of inflated 3D representations onto judgments of the 2D center of mass, we created inflated versions of our shape stimuli (using, without modification or parameter fitting, the original code and parameters from Twarog et al., 2012) and projected the centers of mass of the resulting inflated objects back onto our 2D shapes. This simple “model” (i.e., assuming that subjects were attempting to report the inflated 3D shape’s center of mass rather than the actual 2D shape’s centroid) captured deviations in aggregate shape-wise center-of-mass estimates surprisingly well,  $r(28) = .83$ ,  $p < .001$ , correctly predicting the direction of estimate displacement (along the  $x$  and  $y$  dimensions) in 54 of 60 cases (see Figure 4). It is interesting to note that despite this strong correlation, the inflation-based model also robustly “predicted” stronger biases than we in fact observed because the inflationary center of mass was, on average, 31% farther from the 2D center of mass than were subjects’ actual

estimates,  $t(29) = 3.23$ ,  $p < .01$ ,  $d = .59$  (see the shaded-in discrepancy between the regression line and  $y = x$  in Figure 4). This suggests that subjects were not reporting the 3D center of mass *instead of* the 2D center of mass, but rather that their estimates are *biased by* the 3D center of mass (for effects broken down by subject, see Figure S1 in the online supplementary material). It is this phenomenon that we call “inflationary intrusion”: The mental construction of 3D representations of shapes in the natural course of visual shape segmentation and the subsequent effects of this process on the perception of 2D shape geometry.

**Other data and theories.** Whereas previous work has touched on issues conceptually related to center-of-mass perception, including visual localization (e.g., Cohen & Singh, 2006; Davi, Thomas Doyle, & Proffitt, 1992; Denisova et al., 2006; Vishwanath & Kowler, 2003), judgments of stability (e.g., Battaglia, Hamrick, & Tenenbaum, 2013; Cholewiak, Fleming, & Singh, 2013, 2015; Samuel & Kerzel, 2011; see also Bonawitz, van Schijndel, Friel, & Schulz, 2012), and actual grasping of objects (e.g., Bingham & Muchisky, 1993; Goodale et al., 1994; Iberall et al., 1986; Lederman & Wing, 2003; Lukos, Ansuini, & Santello, 2007), we know of only one previous study focusing on free, explicit center-of-mass estimation (Baud-Bovy & Soechting, 2001). That study found errors in center-of-mass perception for much simpler shapes (e.g., irregular triangles and quadrilaterals) but attributed such errors to an “incenter bias,” according to which an object’s perceived center of mass would be determined by the center of the largest circle that can fit inside of the shape. However, we suggest here that such an apparent incenter bias was actually a special case of a broader inflationary principle of shape processing: If center-of-mass perception is biased toward the inflated shape’s center of mass, then it will also inevitably tend toward the incenter, if only because the center of the largest circle that can fit inside of the shape will also happen to be the thickest region of the inflated shape (e.g., using puffball’s “wherever you can place a circle, place a sphere” rule, which will always place the largest sphere at the incenter).

However, it is important to note that our results suggest that center-of-mass estimates become increasingly decoupled from the incenter for more complex shapes (for an example, see Figure 5). When we directly compared a bias toward the inflated center of mass against a bias toward the incenter, the “inflationary intrusion” hypothesis better tracked subjects’ estimates. For example, compared with an incenter bias, a bias toward the centers of mass of the shapes’ inflated forms was closer to subjects’ actual center-of-mass estimates (25 pixels vs. 47 pixels,  $t(29) = 4.24$ ,  $p < .001$ ,  $d = .78$ ), accounted for more variance in center-of-mass estimation errors (70% vs. 59%;  $z = 2.12$ ,  $p < .05$ ), and nonparametrically outperformed an incenter model, giving predictions that were closer to subjects’ actual estimates in 87% of cases (26 of 30, binomial probability test,  $p < .001$ ). In this way, the new data

Figure 2 (opposite). Results from the center-of-mass estimation task. Green star indicates true center of mass (i.e., the mean of the coordinates of the shape’s interior points), blue circle indicates average location of all subjects’ estimates, and red arrow indicates direction and magnitude of displacement from true center of mass predicted by inflation (the inflated center of mass is located at the arrowhead’s tip). The blue circle on each shape is at least 3 times as large as the standard error for the location it represents (i.e., subjects’ estimated center of mass). Shapes are sorted from top left to bottom right in order of estimation error relative to inflationary center of mass. CoM = center of mass. See the online article for the color version of this figure.

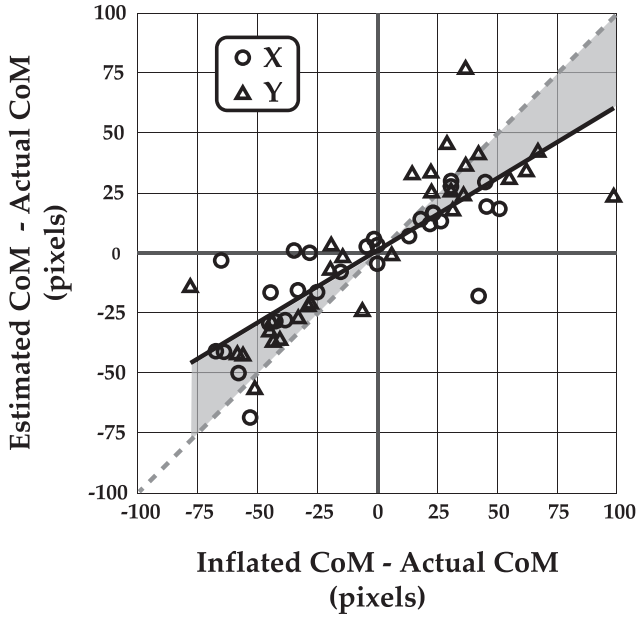


Figure 4. Correlation between estimates' observed deviation from the center of mass and the deviation "predicted" by inflation. Circles represent deviation in the *x* dimension; triangles represent deviation in the *y* dimension. Solid black line is the regression line for all points, and dashed gray line represents a hypothetical 1-to-1 relationship between inflation and center-of-mass estimation. CoM = center of mass.

reported here with more complex shapes cannot easily be reconciled with an incenter bias, whereas the inflationary intrusion account has a ready explanation of previous data that had initially seemed suggestive of an incenter bias.

Ruling out such an incenter bias also helps to rule out other potential interpretations of these results. For example, at least at first glance, an alternative explanation of these biases could be that subjects simply avoided the edges of the images they were asked to click so as to "play it safe" and not accidentally click outside of the shape (and thus fail the task). However, that strategy would lead to exactly the sort of incenter bias that the present data rule out, because the incenter is, by definition, the single "safest" point on the shape to click (being the point with the greatest distance to the nearest edge). Moreover, the strategic avoidance of edges would seem to make odd or even absurd predictions in several cases. For example, the safest place to click the dumbbell shape that served as one of the catch shapes (Figure 2, leftmost catch shape) is the center of one of its lateral "bells" (which is also the location of its incenter), and the "least safe" place to click is the narrow strait at the shape's center. However, that narrow strait is precisely where subjects clicked: The blue circle in that figure not only represents the mean estimate for the dumbbell shape, but it also overlaps with the click locations of 98% of subjects, who apparently did not hesitate to click a location very near the edge (and far away from the incenter).

Thus, a bias toward the incenter provides an inferior fit to the observed data and makes easily falsified predictions for particular cases. These considerations effectively undermine this alternative account and instead support the inflation account of biased center-of-mass perception.

### Study 2: Shapes Versus Nonshapes in Center-of-Mass Perception

We are proposing that visually locating an object's center of mass is biased by an implicit 3D representation of 2D shapes. As such, our account is specific to *objects* and *shapes* per se; thus, it predicts not only the presence of such biases in shape perception (as in Study 1) but also their relative absence for stimuli that are not shapes.

Study 2 tested this strong prediction by running the same task as in Study 1 but with nonshapes substituted as stimuli. In particular, we took the original shapes from Study 1 and derived from them arrays of dots whose boundaries matched those of the original shapes but which had no explicit bounding contour (Figure 6, a and b). The center of mass of a 2D shape is identical to the centroid of each of its pixels, so in principle the visual system should be just as good (or bad) at estimating the centroid of a shape as the centroid of an array of dots arranged in that shape. However, we reasoned that if inflation-based shape segmentation is truly a source of the estimation errors observed in Study 1, then such biases should be reduced for centroid estimates of such dot arrays, which share many geometric properties with the shapes in Study 1 but are not literally shapes in the first place, and so may not be inflated by the visual system. Of course, transforming outlined shapes into dot arrays does not completely eliminate shapeliness because even arrays of dots can be visually segmented (Cohen, Singh, & Maloney, 2008). For this reason we still expected to observe systematic biases consistent with inflation, but we predicted that the average location of each subject's estimate would

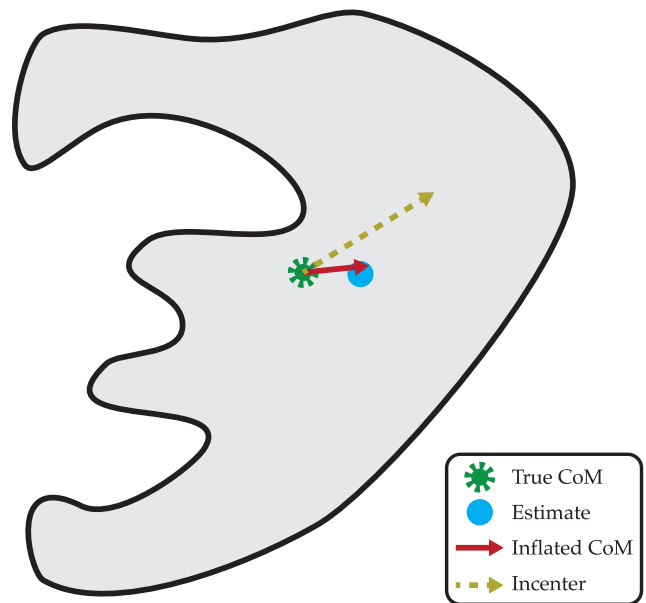


Figure 5. Example comparison of an inflationary bias and an incenter bias. For complex shapes, a bias away from the true center of mass (green star) and toward the inflationary center of mass (solid red arrow) better captures subjects' estimates (blue circle) than does a bias toward the incenter (dashed yellow arrow). For the incen-ters of every shape used in the experiment, see Figure S2 in the online supplemental material. CoM = center of mass. See the online article for the color version of this figure.

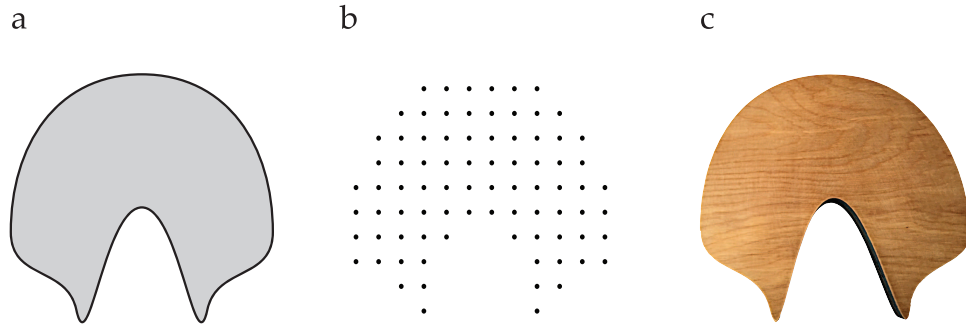


Figure 6. Sample stimuli for Studies 1–3. (a) A 2D shape from Study 1. (b) An array of dots derived from the original 2D shape, as used in Study 2. (c) Photograph of a real-life, uniformly thick object laser-cut out of wood using the original 2D shape as a blueprint (as used in Study 3). See the online article for the color version of this figure.

be more accurate (i.e., closer to the true centroid) for the dot arrays than for the shapes from which they were derived.

## Method

The method was identical to Study 1 except as follows. Two hundred new subjects from Amazon Mechanical-Turk completed the center-clicking task from Study 1, except now subjects were randomly assigned to see either the exact same shapes as in Study 1 (the Shapes condition) or instead to see arrays of dots derived from those shapes (the Dots condition). The dot arrays were created by superimposing a grid of uniformly spaced dots (10 pixels in diameter, with 48 pixels of interdot spacing) on the original shapes; all of those dots falling on the surface of the original shape formed the dot array used as the stimulus. As before, both groups were instructed to click the centroid of the shapes or dots, or “the point on the figure where it would be perfectly balanced.”

## Results and Discussion

**Replicating inflationary intrusion.** For the Shapes condition, center-of-mass judgments showed precisely the same pattern of errors as in Study 1 (and indeed, the location of the subjects’ collective estimates themselves differed by a shape-wise average of only 4.7 pixels between experiments; for a comparison of estimates across experiments, see Figure S2 in the online supplementary material). A bias toward the 2D shapes’ inflated centers of mass was again highly correlated with the observed biases,  $r(28) = .83$ ,  $p < .001$ , and correctly predicted the direction of estimate displacement (along the  $x$  and  $y$  dimensions) in 53 of 60 cases. Thus, we replicated the inflationary intrusion pattern observed in Study 1.

**Dots versus Shapes.** Aggregated center-of-mass estimates were more accurate in the Dots condition than in the Shapes condition. The average proximity (across the 30 matched pairs of stimuli) of the location of the average estimate (across 100 subjects) improved by 16%,  $t(29) = 3.05$ ,  $p < .006$ ,  $d = .55$  (see Figure 7). Thus, subjects were better at finding the centroids of dot arrays than the centers of mass of their equivalent shapes, even though such tasks could be considered computationally identical in principle (because the center of mass of a 2D shape just is the

centroid of its pixels). This suggests, just as our account predicts, that representation of an image *as a shape* drives biased center-of-mass perception.

One potential concern is that the shapes-to-dots transformation inevitably made the resultant images smaller than the original images, and that this could seem to “improve” centroid estimates simply by allowing less room for error. However, there was no correlation between the amount a given shape “shrank” upon being transformed into a dot array and the improvement in centroid estimates for that shape,  $r(28) = .07$ ,  $p > .7$ , suggesting that this simple and small change in size cannot account for the robust improvement in centroid estimates.

Thus, these results support the inflation-based account of biased center-of-mass perception, and they continue to rule out alternative interpretations. For example, if the errors in center-of-mass estimation in these tasks were simply the result of strategic mouse-clicking (e.g., avoiding the shapes’ edges so as not to accidentally click outside of the shape), then there is no reason that altering the

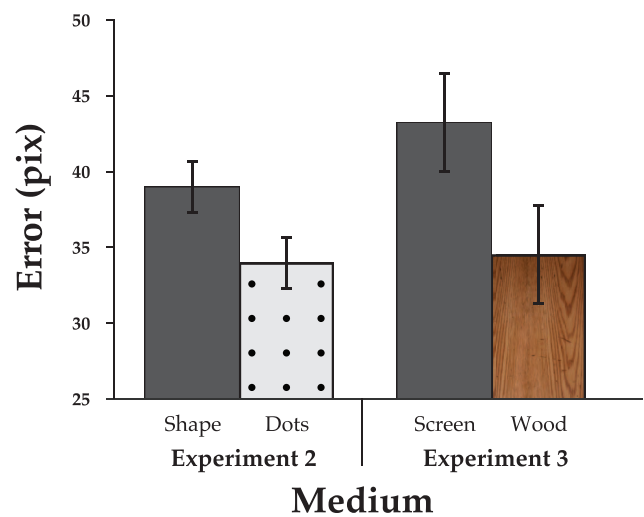


Figure 7. Results from Studies 2 and 3. Center-of-mass estimates were worse for contoured 2D shapes (gray bars) than they were for arrays of dots (Study 2) or real-life uniformly thick blocks of wood (Study 3). Error bars are  $\pm 1$  SEM. See the online article for the color version of this figure.

images' shapehood should have any impact on subjects' estimates. That such a manipulation did alter—and in fact improved—such estimates suggests that shape representation per se plays a key role in the genesis of these effects.

### Study 3: Real-Life 3D Objects

We have suggested that, in processing 2D shapes, the mind constructs 3D representations of such shapes, and so it is natural that the foregoing experiments have used 2D images as stimuli. However, one context in which the visual system may be less likely to rely on such 3D representations (and perhaps less likely to be misled by such representations) is when the stimulus already contains explicit information about the shape's depth in various regions—for example, when the stimulus is a real-life 3D object whose thickness can be directly appreciated. Indeed, the original impetus for inflationary approaches to shape segmentation is that segmentation of volumetric 3D objects is more reliable than segmentation of 2D silhouettes. By this logic, it may be less necessary for the visual system to infer the inflated 3D form of a shape when the shape's actual 3D structure is readily accessible.

Thus, Study 3 used such real-life 3D objects as stimuli. We compared center-of-mass judgments for the shapes used in Study 1 against center-of-mass judgments for real-life, uniformly thick versions of those shapes (custom-cut out of wood using a commercial laser). We reasoned that if inferences to an inflated, nonuniformly thick, 3D form of the original 2D shapes contributes to errors in center-of-mass perception, then directly confronting subjects with the shape's true 3D structure should reduce such biases and result in more accurate center-of-mass estimates.

### Method

**Subjects.** Twenty Yale University students participated in exchange for course credit.

**Stimuli.** We used a 150-W Universal Laser Systems ILS 12.150D laser cutter to construct, out of birch plywood, uniformly thick (1.3 cm) objects, the shapes of which matched those from Study 1. The objects were sanded and finished with a moisture-resistant polyurethane coating.

**Procedure.** Each subject completed two tasks (with the order of the tasks counterbalanced between subjects). In the Screen task, subjects completed the same estimation task as in Study 1 for a subset of 15 of the 30 experimental shapes (sampled randomly and independently for each subject). In the Wood task, the experimenter placed in front of the subject one of the wooden objects and asked the subject to estimate its center of mass—by sight only, never physically handling the object—and then to gently place his or her left index finger on the surface of the object to mark the estimate's location. Then, on a screen in front of the subject, the 2D shape image from Study 1 matching the wooden object's shape appeared, and subjects used their right hand to click the point on the on-screen shape corresponding to the point on the wooden shape where their left index finger was currently located. Subjects were encouraged to stick with their original estimate and not to change their mind during the “transfer” of their estimate from the wooden object to the on-screen shape. This task was repeated for whichever 15 shapes were not selected for the Screen task. Thus, subjects completed an identical task at the time their responses

were recorded, with the only difference between conditions being whether the subjects first saw a uniformly thick, real-life version of the shape.

### Results and Discussion

Collapsing across both task orders (Wood-Screen and Screen-Wood), subjects' responses trended in favor of more accurate estimates for the wooden objects over the on-screen shapes, the latter of which were 9% farther from the true center of mass than the former,  $t(29) = 1.15$ , *ns*. However, we suspected (and several subjects themselves suggested in debriefing) that task order played a role in performance, such that, for example, initially seeing a set of shapes in wood would subsequently alter subjects' understanding of the on-screen shapes (e.g., by leading them to assume that the on-screen shapes were also uniformly thick and to consequently correct for this). Indeed, considering only the first task completed by each subject (i.e., Wood for half of the subjects and Screen for the other half), in which no such contamination could have occurred, there was a reliable improvement in center-of-mass estimates for the wooden shapes over the on-screen shapes, with estimates for on-screen shapes falling 25% farther from the true center of mass than for wooden shapes,  $t(29) = 2.64$ ,  $p = .01$ ,  $d = .48$  (see Figure 7)<sup>4</sup>. (This effect was robust to a conservative Bonferroni correction [ $\alpha = .025$ ] for multiple comparisons.) Moreover, the improvement in center-of-mass estimation appeared to benefit from a reduction in the inflation bias in particular: Whereas estimates for the on-screen shapes were closer to the inflationary center of mass than the true center of mass (30 pixels away vs. 43 pixels away,  $t(29) = 2.32$ ,  $p < .03$ ,  $d = 0.42$ ), estimates for the wooden shapes were not (34 pixels vs. 34 pixels,  $t(29) = 0.14$ , *ns*), and the interaction—which captures the tendency of estimates to move away from the inflationary center of mass and toward the true center of mass—was also reliable,  $t(29) = 2.14$ ,  $p < .05$ ,  $d = .39$ . Thus, seeing uniformly thick 3D versions of the shapes improved estimations of their center-of-mass compared with seeing their 2D outlines, and in the way expected if 2D center-of-mass estimation is biased by an assumption of nonuniform thickness.

### General Discussion

In a brief glance at an object or shape, we can appreciate not only its size, color, and location but also a rich suite of functional properties, including the organization of its parts (DeWinter & Wagemans, 2006; Hoffman & Richards, 1984); its underlying body plan (Feldman & Singh, 2006; Firestone & Scholl, 2014; Treder, 2010); its graspability (Lukos et al., 2007); and, apparently, its center of mass. However, both in the real world and in the laboratory, our perception of balance shows puzzling systematic biases. In three studies, we found the first psychophysical evidence that the computer-graphics technique of shape inflation

<sup>4</sup> The second-task data for each subject also allowed us to rule out independent group-level effects. For example, the two groups of subjects had nearly identical performance in the Wood task, averaging 34 pixels (Wood-first group) and 33 pixels (Screen-first group) from the true centers of mass of the wooden shapes,  $t(29) = 0.28$ ,  $p > .75$ , suggesting that it was not simply that one group was better at the task than the other.



may be implemented in the mind, showing that inflation leads center-of-mass perception astray by implicitly assuming that outlined shapes are nonuniformly thick. We found that center-of-mass judgments for 2D shapes are biased in a way uniquely predicted by inflation of those shapes (e.g., as opposed to an incenter bias or strategic estimation), and that such biases are reduced by attenuating and by enhancing the available shape information in the stimuli: Reducing shapehood by converting contoured 2D shapes into arrays of dots improved subjects' ability to perceive their centers of mass, as did explicitly displaying a uniformly thick third dimension (by using real-life blocks of wood). Together, these results suggest that shape inflation is psychologically real and that it intrudes on visual representation of the objects we see.

### From Machines to Minds

Although promising in principle as an approach to shape segmentation, the notion of shape inflation has so far been confined to computer vision models and graphics applications (e.g., transferring textures from one shape to another; Twarog et al., 2012). However, its success in various applied contexts raises the intriguing possibility that this technique is actually implemented in the mind. The results presented here suggest that the visual system does in fact use an inflationary process in representing shapes. Thus, these results provide a case study of how theoretical insights about how perception *could* work in principle can generate testable hypotheses for how perception *does* in fact work in the mind.

### A “Feature,” Not a “Bug”

The principles underlying visual perception are often described as a necessarily imperfect “bag of tricks” (Ramachandran & Anstis, 1986) that can be easily “misled” (as in all visual illusions), but it is noteworthy that a certain finesse is often required to produce illusory visual phenomena. (For example, consider how many visual illusions require accidental views of carefully arranged stimuli, false or unusual lighting conditions, unnatural computer-graphics displays, etc.). By contrast, the biases reported in the present studies are generated with striking ease: There is nothing particularly unusual about the stimuli used in the present experiments, which are simply shapes with salient part structures, and yet they bias our perception of balance in surprisingly large and systematic ways. The ease with which such judgments are led astray in the present cases suggests a deeper, underlying *reason* for such biases—the possibility that these biases reflect an adaptive “feature,” rather than a flawed “bug.” That is precisely the account we have proposed and tested here: That biased perception of balance in such cases reflects a more general functional characteristic of object processing that serves us well in navigating and manipulating a 3D world.

### References

- Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, *19*, 392–398. <http://dx.doi.org/10.1111/j.1467-9280.2008.02098.x>
- Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *PNAS: Proceedings of the National Academy of Sciences of the United States of America*, *110*, 18327–18332. <http://dx.doi.org/10.1073/pnas.1306572110>
- Baud-Bovy, G., & Soechting, J. (2001). Visual localization of the center of mass of compact, asymmetric, two-dimensional shapes. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 692–706.
- Bingham, G. P., & Muchisky, M. M. (1993). Center of mass perception and inertial frames of reference. *Perception & Psychophysics*, *54*, 617–632. <http://dx.doi.org/10.3758/BF03211785>
- Bonawitz, E. B., van Schijndel, T. J. P., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, *64*, 215–234. <http://dx.doi.org/10.1016/j.cogpsych.2011.12.002>
- Cholewiak, S. A., Fleming, R. W., & Singh, M. (2013). Visual perception of the physical stability of asymmetric three-dimensional objects. *Journal of Vision*, *13*, 12. <http://dx.doi.org/10.1167/13.4.12>
- Cholewiak, S. A., Fleming, R. W., & Singh, M. (2015). Perception of physical stability and center of mass of 3-D objects. *Journal of Vision*, *15*, 13. <http://dx.doi.org/10.1167/15.2.13>
- Cohen, E. H., & Singh, M. (2006). Perceived orientation of complex shape reflects graded part decomposition. *Journal of Vision*, *6*, 805–821. <http://dx.doi.org/10.1167/6.8.4>
- Cohen, E. H., Singh, M., & Maloney, L. T. (2008). Perceptual segmentation and the perceived orientation of dot clusters: The role of robust statistics. *Journal of Vision*, *8*, 1–13. <http://dx.doi.org/10.1167/8.7.6>
- Cole, F., Sanik, K., DeCarlo, D., Finkelstein, A., Funkhouser, T., Rusinkiewicz, S., & Singh, M. (2009). How well do line drawings depict shape? *ACM Transactions on Graphics (TOG): Proceedings of ACM SIGGRAPH 2009*, *28*, 28.
- Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, *8*, e57410. <http://dx.doi.org/10.1371/journal.pone.0057410>
- Davi, M., Thomas Doyle, M. A., & Proffitt, D. R. (1992). The role of symmetry in determining perceived centers within shapes. *Perception & Psychophysics*, *52*, 151–160. <http://dx.doi.org/10.3758/BF03206768>
- Denisova, K., Singh, M., & Kowler, E. (2006). The role of part structure in the perceptual localization of a shape. *Perception*, *35*, 1073–1087.
- De Winter, J., & Wagemans, J. (2006). Segmentation of object outlines into parts: A large-scale integrative study. *Cognition*, *99*, 275–325. <http://dx.doi.org/10.1016/j.cognition.2005.03.004>
- Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences*, *103*, 18014–18019.
- Firestone, C., & Scholl, B. J. (2014). “Please tap the shape, anywhere you like”: Shape skeletons in human vision revealed by an exceedingly simple measure. *Psychological Science*, *25*, 377–386. <http://dx.doi.org/10.1177/0956797613507584>
- Galton, F. (1907). Vox populi. *Nature*, *75*, 450–451. <http://dx.doi.org/10.1038/075450a0>
- Goodale, M. A., Meenan, J. P., Bühlhoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*, *4*, 604–610. [http://dx.doi.org/10.1016/S0960-9822\(00\)00132-9](http://dx.doi.org/10.1016/S0960-9822(00)00132-9)
- Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65–96. [http://dx.doi.org/10.1016/0010-0277\(84\)90022-2](http://dx.doi.org/10.1016/0010-0277(84)90022-2)
- Hoffman, D. D., & Singh, M. (1997). Saliency of visual parts. *Cognition*, *63*, 29–78. [http://dx.doi.org/10.1016/S0010-0277\(96\)00791-3](http://dx.doi.org/10.1016/S0010-0277(96)00791-3)
- Iberall, T., Bingham, G. P., & Arbib, M. A. (1986). Opposition space as a structuring concept for the analysis of skilled hand movements. In H. Heuer & C. Fromm (Eds.), *Generation and modulation of action patterns* (pp. 158–173). Berlin, Germany: Springer-Verlag. [http://dx.doi.org/10.1007/978-3-642-71476-4\\_12](http://dx.doi.org/10.1007/978-3-642-71476-4_12)
- Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. (1992). Surface perception in pictures. *Perception & Psychophysics*, *52*, 487–496. <http://dx.doi.org/10.3758/BF03206710>

- Kovács, I., Fehér, A., & Julesz, B. (1998). Medial-point description of shape: A representation for action coding and its psychophysical correlates. *Vision Research*, 38, 2323–2333. [http://dx.doi.org/10.1016/S0042-6989\(97\)00321-0](http://dx.doi.org/10.1016/S0042-6989(97)00321-0)
- Kovács, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature*, 370, 644–646. <http://dx.doi.org/10.1038/370644a0>
- Lederman, S. J., & Wing, A. M. (2003). Perceptual judgement, grasp point selection and object symmetry. *Experimental Brain Research*, 152, 156–165. <http://dx.doi.org/10.1007/s00221-003-1522-5>
- Lukos, J., Ansuini, C., & Santello, M. (2007). Choice of contact points during multidigit grasping: Effect of predictability of object center of mass location. *The Journal of Neuroscience*, 27, 3894–3903. <http://dx.doi.org/10.1523/JNEUROSCI.4693-06.2007>
- Ramachandran, V. S., & Anstis, S. M. (1986). The perception of apparent motion. *Scientific American*, 254, 102–109. <http://dx.doi.org/10.1038/scientificamerican0686-102>
- Samuel, F., & Kerzel, D. (2011). Is this object balanced or unbalanced? Judgments are on the safe side. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 529–538. <http://dx.doi.org/10.1037/a0018732>
- Treder, M. S. (2010). Behind the looking-glass: A review on human symmetry perception. *Symmetry*, 2, 1510–1543. <http://dx.doi.org/10.3390/sym2031510>
- Twarog, N. R., Tappen, M. F., & Adelson, E. H. (2012). Playing with Puffball: Simple scale-invariant inflation for use in vision and graphics. In *Proceedings of the 2012 ACM Symposium on Applied Perception* (pp. 47–54). New York, NY: Association for Computing Machinery.
- Vishwanath, D., & Kowler, E. (2003). Localization of shapes: Eye movements and perception compared. *Vision Research*, 43, 1637–1653. [http://dx.doi.org/10.1016/S0042-6989\(03\)00168-8](http://dx.doi.org/10.1016/S0042-6989(03)00168-8)
- Wilder, J., Feldman, J., & Singh, M. (2011). Superordinate shape classification using natural shape statistics. *Cognition*, 119, 325–340. <http://dx.doi.org/10.1016/j.cognition.2011.01.009>

Received May 23, 2015

Revision received December 25, 2015

Accepted January 11, 2016 ■

### Correction to Foroughi et al. (2015)

In the article “Interruptions Disrupt Reading Comprehension,” by Cyrus K. Foroughi, Nicole E. Werner, Daniela Barragán, and Deborah A. Boehm-Davis (*Journal of Experimental Psychology: General*, 2015, Vol. 144, No. 3, pp. 704–709. <http://dx.doi.org/10.1037/xge0000074>), the effect sizes (Cohen’s  $d$ ) reported used the following formula:

$$d = \frac{t}{\sqrt{n}}$$

The authors later found out that this formula should not be used to calculate Cohen’s  $d$  for a paired sample (Cortina, personal communication). Therefore, they present the correct effect sizes using the traditional formula for Cohen’s  $d$  (see Dunlap, Cortina, Vaslow, & Burke, 1996 for a review):

$$d = \frac{M_2 - M_1}{SD_{pooled}}$$

The correct effect sizes are:

Experiment 1:  $d = .85$ ; Experiment 2:  $d = .95$ ; Experiment 3:  $d = .70$ ,  $d = .73$

Notably, the effect sizes are still considered “large” (or near large) by conventional standards (Cohen, 1998), and this error has no bearing on the implications of the research.

### References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1(2), 170–177. <http://dx.doi.org/10.1037/1082-989X.1.2.170>

<http://dx.doi.org/10.1037/xge0000186>