Volume Estimation through Mental Simulation

Hannah Perfecto & Clayton R. Critcher

University of California, Berkeley

Abstract

Although mental simulation underlies many day-to-day judgments, we identify a new domain influenced by simulation: volume estimation. Previous research has identified various ways that volume estimates are biased, but typically has not presented a psychological process by which such judgments are made. Our *simulation-informs-perception* (SIP) account proposes people often estimate a container's size by simulating filling it. First, this produces an orientation effect: The same container appears larger right-side-up than upside-down, due to the greater ease of imagining filling an upright container. Second, we identify a cavern effect: Imagining pouring through a narrow opening into a relatively wide base produces a sense the container is cavernous (compared to identically-sized wide-topped, narrow-based containers). By testing for and demonstrating the importance of the cognitive process of simulation to these effects, we show how complex perceptual judgments can be distorted by top-down influences even when they are necessarily informed by modularly-processed perceptual input.

KEYWORDS: simulation, judgment, perception, volume estimation

Volume Estimation through Mental Simulation

Humans' capacity for mental simulation assists with a variety of judgments. It helps forecast the future: Grocery shoppers simulate how hungry they will be when determining how much to buy (Gilbert, Gill, & Wilson, 2002). Global warming seems more likely to occur when it is easier to simulate its effects (Risen & Critcher, 2011). Although simulation can offer valid input to judgments, simulations can also be distorted by uninformative states. Hungry shoppers wrongly project their present hunger into the future and overbuy. Those in cold rooms have trouble simulating the consequences of a world plagued by global warming and thus err toward climate change skepticism. Given people frequently mistake changes in the self for changes in the world (Eibach, Libby, & Gilovich, 2003), it is perhaps unsurprising that people fail to correct for these biases.

More controversially, however, researchers have claimed that mental simulation affects visual perception of the present. People estimate their distance from an object by simulating reaching for it (Linkenauger, Witt, Stefanucci, Bakdash, & Proffitt, 2009). When assessing the steepness of a hill, people imagine climbing it (Witt & Proffitt, 2008). Those who can jump higher may envision more reachable and thus shorter world. Those wearing heavy backpacks or who have a fear of falling imagine a climb to be more tiring and treacherous than those less burdened and more energized, and accordingly report slopes to be steeper and shallower, respectively (Proffitt, 2006; Proffitt, Stefanucci, Banton, & Epstein, 2003; Schnall, Zadra, & Proffitt, 2010; Stefanucci, Proffitt, Clore, & Parekh, 2008).

Although simulation is necessary to consider and judge the future, the idea that cognitive processes can exert a top-down influence on visual perception has come under recent scrutiny (Firestone, 2013; Firestone & Scholl, in press). After all, a hallmark of visual processing is modularity, meaning it should be immune to cognitive interference (Fodor, 1983 Pylyshyn, 1999). Indeed, Firestone and Scholl (2014) show that many supposedly top-down effects on perception are ultimately attributable to artifactual characteristics (see also Durgin, Klein, Spiegel, Strawser, & Williams, 2012).

In the present paper, we do not claim to settle this debate, but rather demonstrate how cognition may play a crucial (and distorting) role in translating perceptual input into more complex perceptual judgments. Consider volume perception. Although we *perceive* size, the width is less direct than the simple single-dimensional perceptions that inform it (e.g., height). Instead, volume perception requires people to synthesize or translate perceptual inputs into a perceptual judgment. Of course, this can be done formally (i.e., with a ruler, a geometry textbook, and a calculator). In practice, people approach such perceptions more intuitively; one rarely observes—for example—coffee shop patrons leaning on formal instrumentation or squared radii when choosing a cup size. Previous research has identified how salient features (e.g., height) exert disproportionate influence in volume perception (Wansink & van Ittersum, 2003; Yang & Raghubir, 2008). However, this literature largely provides “as-if models” (Ordabayeva & Chandon, 2013)—algorithms that can anticipate specific sources of judgment error, but that remain agnostic about the process by which those perceptual judgments are made.

We propose that mental simulation is a cognitive process that often converts modularly-processed, perceptual input into the perceptual judgment of volume. To estimate a receptacle’s volume, people often simulate how much they could pour into it. Pouring happens with the flow of gravity (from top to bottom) into an upright container. Our *simulation-informs-perception* (SIP) account leans on this simple property to deduce two novel hypotheses of how simulation may lead volume perception astray.

First, we hypothesize an *orientation effect*: A container will seem larger when right-side-up than upside-down. We suggest this because it is easier to imagine filling a right-side-up container than an upside-down one, and that the ease of this simulation contributes positively to the subjective perception of its size. By analogy, if it is hard to simulate fitting one's furniture in a room, that room may seem smaller.

Second, we posit a *cavern effect*. Because one fills a container by pouring through its opening to reach its base, a low top-to-base ratio will create a sense that a container is large and cavernous. That is, imagining liquid descending through a narrow top toward a relatively wide, open base may present a stark contrast that offers the subjective sense of filling a vast space. Our studies test for both effects, determine whether these effects are driven by simulation (instead of responses to low-level differences in targets' features), assess whether subjective ease of simulation produces the orientation effect, and pinpoint whether it is indeed a low opening-to-base ratio (instead of the container's shape) that produces the cavern effect.

Experiment 1

Experiment 1 tested the orientation effect—whether the same cup looks larger when right-side-up than upside-down.

Method

Participants and design. In an effort to achieve large sample sizes, in all of our experiments, we recruited participants simultaneously from an undergraduate subject pool (online: Experiment 1; lab: Experiments 2-4) at the University of California, Berkeley, as well as Amazon's Mechanical Turk. Research assistants, who were scheduled to work in the lab a certain number of hours each week, recruited as many participants as they could over a specified amount of time—typically, until the end of a semester. The sample size on Mechanical Turk was determined by: 1) the total budgeted funds for Mechanical Turk for a particular month, and 2) how many other experiments the lab was running on Mechanical Turk in that month. That these guidelines allowed us to far exceed the recommended minimum sample size of 50 participants per cell (Simmons, Nelson, & Simonsohn, 2013), combined with our within-subjects designs, led us to believe our sample sizes were appropriate. Although we discuss all manipulations, exclusions, and hypothesis-relevant measures in the main text, we invite interested readers to consult the Supplemental Materials for discussion of three additional measures (one exploring participants' memory for stimuli [Studies 1-4] and two assessing self-reported compliance with instructions [Study 4].) The full materials and data for each study are available here: https://osf.io/f67yr/?view_only=68a5f0aff1c342049ba5934307c626df.

Two hundred eight-two participants took part in Experiment 1. On some trials, participants indicated the volume of a cup was 0 ounces (52 of 6,226 responses; 0.84%). In 51 of 52 cases, participants were judging an upside-down cup; thus, some participants seemed to think these were trick questions ("An upside-down cup cannot hold any liquid!"). Because inclusion of these trials would artifactually promote hypothesis-consistent results, we excluded these trials (in this and all experiments).

Materials. Participants were asked to estimate the volume of 24 cups. In order to attain a diversity of stimuli, we sampled cups from Google Images. We began by locating images of six, distinct cups. In particular, we found two wide-based cups (e.g., a stemless wine glass), two wide-topped cups (e.g., a standard water glass), and two cylindrical containers (e.g., a coffee mug). We found a red (opaque) and transparent (clear) cup of each shape. For each of the six cups, we created a large and a small version. The larger version simply expanded the width and height of the small version by sqrt(2) so that the surface area covered by the image was exactly twice as big for the small version as the large version. Note this makes the implied volume of the large image 2.82 [i.e., sqrt(2) ^ 3] times the volume of the corresponding small cup.

This gave us twelve unique cups: a small and large version of the six original images. Most central for our purposes, we then created an upside-down and a right-side-up version of each cup. The upside-down version was the same image as the right-side-up version, except it had been rotated 180 degrees. Figure 1 depicts example stimuli.

Figure 1. Sample stimuli for Experiment 1: wide-topped (left column), cylindrical (middle column), and wide-based (right column) clear (top row) and opaque (bottom row) cups.

Procedure. Both to familiarize all participants with the unit used (ounces) and because all images were presented on a computer screen, we first provided participants with a modulus that helped participants get a sense of scale (Raghubir & Krishna, 1999). Before each trial, participants saw an image of the same generic soda can. Participants were informed that the can was exactly 12 ounces. This permitted us to more confidently conclude that systematic differences in perceptions of volume would be explained by differences in perception (e.g., a judgment of 15 ounces reflects a perception that a cup is 25% larger than the preceding cup) instead of different understandings of the unit of judgment (e.g., "What number of ounces would match my *subjective* perception of 'pretty big?'"). Then participants judged the perceived volume of the container on a slider scale that ranged from 0 ounces to 24 ounces. The slider defaulted to 12 ounces (the size of the modulus) and permitted responses in tenth-ounce increments. Participants made these judgments for all 24 cups in a randomized order.

Results

Because some participants had missing data (either because they indicated a cup's size was 0 ounces, they skipped an item, or exited the experiment early), we opted to use multi-level modeling instead of a repeated-measures ANOVA. This also establishes consistency in the analytic technique used across our experiments given certain later experiments cannot be analyzed with a repeated-measures ANOVA.

We defined one Level-1 variable: *orientation*. This differentiated trials for which a cup was depicted right-side-up (+1) or upside-down (-1). We nested orientation within participant in a random-slope, random-intercept model. This permitted the effect of orientation to vary for each participant (random-slope) and accounted for differences between participants in how much they tended to see targets as larger vs. smaller (random-intercept). We also created a categorical variable *cup*, which identified each of the 12 unique cups depicted. In this and all subsequent studies, an identically-named variable is used to identify stimuli that vary only on the key factors of interest (e.g., orientation). This variable was included as a random effect, which permitted us to account for variation in volume judgments that was attributable to factors that were not of interest (e.g., size, color).

Consistent with the hypothesized orientation effect, we observed a main effect of orientation, $t(5989.90) = 2.90$, $p = .004$. This reflected that participants saw they exact same cup as larger when it was depicted right-side-up (*M* = 11.43 ounces) as opposed to upside-down (*M* = 11.21 ounces). In other words, turning a cup from upside-down to right-side-up increased its perceived volume by 1.96%. Although one could certainly view this boost as modest, we note both that this is the smallest of the effect sizes we observe and that this robust effect was achieved through extremely modest means—by merely turning an image on its head.

Experiment 2

Experiment 2 was designed to test both the orientation and the cavern hypotheses. The wide-topped and wide-based cups in Experiment 1 were not matched on their shape and size, so we were not able to test the cavern hypothesis. Experiments 2-4 sidestepped this limitation by using digitally-generated containers, which allowed us to hold all properties of a stimulus constant except for its orientation and top-to-base ratio. We predicted we would find that right-side-up containers would seem bigger than equivalent upside-down containers (the orientation hypothesis) and that containers with small top-to-base ratios would appear larger than identically-sized receptacles with large top-to-base ratios (the cavern hypothesis).

Beyond merely demonstrating these two effects, we wanted to determine whether they do indeed result from simulation. Our logic was premised on effects that would emerge as people imagined filling an empty vessel. For half of participants, we explicitly asked them to make the volume judgments by imagining how much they could pour into the container. In this *fill* condition, we expected to find support for both the orientation and cavern hypotheses. The other half of participants were told the container was currently full to the brim (a claim reinforced by the depiction of the cup). They were asked to imagine how much would be poured out if the cup were completely emptied. In this *empty* condition, the conditions that were supposed to give rise to our two effects (the ease of imagining pouring into an upright cup through a narrow opening into a wide base) would not hold. As such, we expected the orientation and cavern effects would be significantly reduced or eliminated. Such a pattern would not only demonstrate the importance of simulation to our effects, but would show that they do not simply reflect a perceptual illusion stemming from the low-level visual features of the stimuli themselves.

Method

Participants and design. Two hundred fifty participants took part in Experiment 2. Participants were randomly assigned to one of two *simulation* conditions: fill or empty. We excluded 7 of the 6,006 (0.12%) responses because the participant made an estimate of zero volume. In all 7 cases, the target cup was upside-down and the participant was in the *empty* condition. In other words, these participants seemed to be indicating that nothing could be held inside an upside-down cup; hence, nothing could be poured out.

Materials. Once again, participants were asked to estimate the volume of 24 digitally-generated cups. These stimuli were created by generating containers that reflected every combination of four factors. Three of these factors defined the cups themselves: color (blue or green), size (small or large), and orientation (wide-based, wide-topped, cylindrical). See Figure 2 for sample stimuli. The volume of the large cup was approximately 2.7 times that of the small cup.

Wide-based and wide-topped cups were truncated cones. These cups were identical in shape, but varied only in whether the larger end of the truncated cone was depicted as the base (wide-based) or the top (wide-topped). We communicated which side of the receptacle was the top of the cup, and thus the container's orientation, by an affixed label that read "Crystal Lakes." That is, whether the label appeared right-side-up or upside-down revealed that the cup's orientation and, by extension, whether it had a low or high top-to-base ratio.

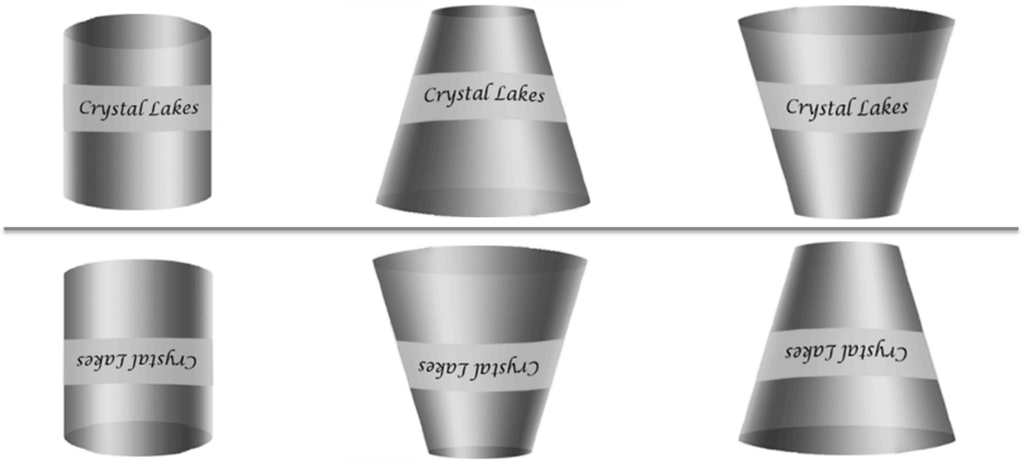

Figure 2. Sample stimuli for Experiment 2: right-side-up (top row) or upside-down (bottom row) cylindrical (left column), wide-based (middle column), or wide-topped (right column) cups.

Procedure. Like in Experiment 1, each trial began with the presentation of a modulus to remind participants of the scale. Although we used the familiar metric of fluid ounces in Experiment 1, we introduced a fictitious unit of measurement for the remaining studies: xids. Familiarity with typical cup sizes in ounces might have constrained at least some participants' willingness to report their actual perception of nuanced differences in size. The modulus container was always said to be 15 xids. On each trial, participants adjusted from 15 xids along a slider scale to arrive at their final volume estimates. The allowable response range was 0 to 30 xids, in tenth-xid increments. In the *fill* condition, participants were asked to "imagine filling this empty cup all the way to the brim" In the *empty* condition, participants were instead asked to imagine pouring all of the liquid out of the cup, "which is currently filled all the way to the brim."

Results

We followed a very similar analytic approach to that taken in Experiment 1. First, we defined two Level-1 variables: *orientation* and *shape*. Orientation differentiated cups that were depicted right-side-up (+1) vs. upside-down (-1). Shape instead differentiated cups that were wide-based (+1), cylindrical (0), or wide-topped (-1). Both shape and orientation were nested within participant in a random-slope, random-intercept model. This permitted the effect of each factor to vary by participant (random-slope) and accounted for differences between participants in how they used the xid scale (random-intercept).

We defined the Level-2 variable *simulation*, which distinguished those participants who were asked to estimate volume through a mental simulation of *fill*ing a cup (+1) from those who were asked how much they could *empty* from the cup (-1). We included the three two-way interactions and one three-way interaction made from crossing the orientation, shape, and simulation variables. Finally, we again included the random effect of *cup* to account for theoretically-irrelevant variance attributable to other bottom-up features. In testing for the orientation and cavern effects, we tested a single model, but attend to different terms; furthermore, we test the cavern effect using a subsample of the stimuli.

Orientation effect. We began by assessing whether we replicated the orientation effect, at least when participants were encouraged to estimate volume by simulating filling a container (as opposed to emptying it). Consistent with the SIP account, the Simulation X Orientation interaction was significant, $t(6197.00) = 4.68$, $p < .001$ (see Figure 3, top panel). When participants were encouraged to estimate volume by imagining filling up the cup, the exact same cup appeared to be larger when right-side-up (M = 18.44 xids) as opposed to upside-down (M = 17.60 xids), $t(6194.21) = 6.70$, $p < .001$. But when volume was estimated by a different simulation—by imagining emptying the cup—the effect of orientation disappeared (*M*s = 17.80 xids vs. 17.76 xids), $t(6199.10) = .33$, $p = .742$.

Cavern effect. In order to assess the cavern hypothesis, we restricted our analysis to the 16 (of 24) cups that were wide-topped or wide-based. This permitted us to analyze participants' estimates of the 16 truncated cones, all of which were one of two sizes (given all small and all large cups were actually identically shaped). It is worth reinforcing that, because we varied the orientation and the shape orthogonally, the specific shape depicted on the screen took only one of two forms. It was the inclusion of the brand label that allowed us to change the conceptual understanding of the container's shape—whether it would be filled by pouring through a narrow top into a wide base (a low top-to-base ratio that should underlie the cavernous illusion that defines the cavern effect), or whether it would be filled by pouring through a wide top into a narrow base.

Once again and consistent with our SIP account, we found that the cavern effect depended on the nature of the simulation participants engaged in (see Figure 3, bottom panel). More specifically, we found a strong Simulation X Shape interaction *t*(4864.20) = 5.70, *p* < .001. As hypothesized, when participants were imagining filling up the cup, the same truncated cone appeared larger when the narrow and wide ends were depicted as being the top and base, respectively (*M* = 19.46 xids) than when this pairing was reversed (*M* = 17.62 xids), *t*(4874.42) = 10.95, *p* < .001. But when the volume was estimated by empting a cup, the cavern effect was reduced (*M*s = 18.45 xids vs. 17.90 xids), *t*(4847.72) = 3.41, *p* = .001.

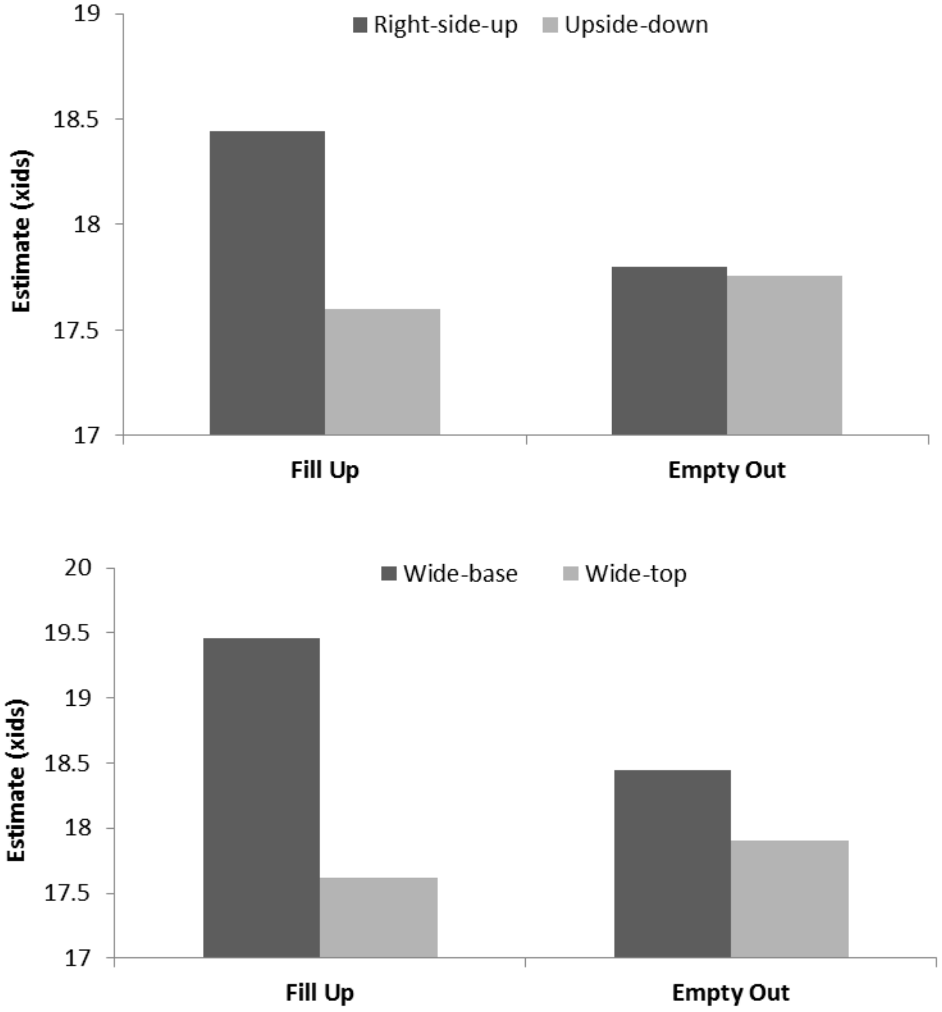

Figure 3. Moderation of the orientation effect (top panel) and the cavern effect (bottom panel) by simulation instructions (Experiment 2). The SIP account predicts orientation and cavern effects to be stronger in the fill condition.

Although we used a relatively minimal manipulation (the orientation of a brand label) to alter the cups' orientation and top-to-base ratio without changing much in the way of the bottom-up details of the shape being judged, it is fair to ask whether these effects reflect a distortion in the interpretation of the image (e.g., Might the brand label's orientation change the perceived curvature of the sides? Might the partially occluded wide circle be assumed to be larger when it represented a container's base instead of its top?) Such possibilities—although we had no reason to suspect them—are different from our SIP account that simulation informs complex size perceptions. Yet such concerns highlight the clear value of the simulation manipulation. Even with exposure to the same low-level features, those instructed to estimate the volume through the simulation of filling (as opposed to emptying) showed stronger orientation and cavern effects. Given that participants in Experiment 1 were not directed to estimate cups' volume though any particular simulation, our results are consistent with our account that (many) people spontaneously approach volume estimation as a simulated judgment task. (Experiment A, reported in the Supplemental Materials, shows, among other things, that the cavern effect also emerges without any reference to simulating filling the containers, consistent with the idea that people lean on simulation *spontaneously* to inform volume estimates.)

Experiment 3

Experiment 3 explores why the orientation and cavern effects emerge. First, we directly tested whether it is the greater ease of simulating filling a right-side-up container (vs. an upside-down one) that underlies the orientation effect. Second, we further probe the cavern effect by varying the top-to-base ratio to determine whether the cavern effect is strongest when the base and top are most unequal in size.

Method

Participants and design. One hundred eighty-nine participants took part in Experiment 3. Eight of 4,998 trials (0.16%) were excluded given estimates of zero xids.

Materials. As in Experiment 2, participants made judgments about 24 computer-generated cups that varied in their orientation, shape, and size. In this experiment, all cups were the same color: blue. Unlike in our previous experiments, all shapes were truncated cones; no cylinders were included. Furthermore, we varied the size of the small and large ends to produce larger or smaller discrepancies between the width of the opening and the base. This permitted us to test the key role of the top-to-base ratio in the cavern effect. Also, we added dotted lines to represent occluded sections of the cups' bases (see Figure 4). Participants were reminded of the meaning of the dotted lines on every trial.

For high-disparity cups, the diameters of the top and base were quite disparate: 4 units and 9 units, respectively (wide-based), or 9 units and 4 units, respectively (wide-topped). For low-disparity cups, the diameters of the top and base were more similar: 4 units and 6 units (or the reverse), or 6 units and 9 units (or the reverse). In addition to using tops and bases characterized by the same ratio, both low-disparity cups shared one end in common (either the top or the base) with the high-disparity cup. We modified the height of each container so that all large cups had the same volume; all small cups had the same volume as well. This means the height of the two low-disparity cups straddled the height of the high-disparity cup. We review all of these details to demonstrate the care that was taken to minimize the possibility that confounding features could explain the predicted pattern of results.

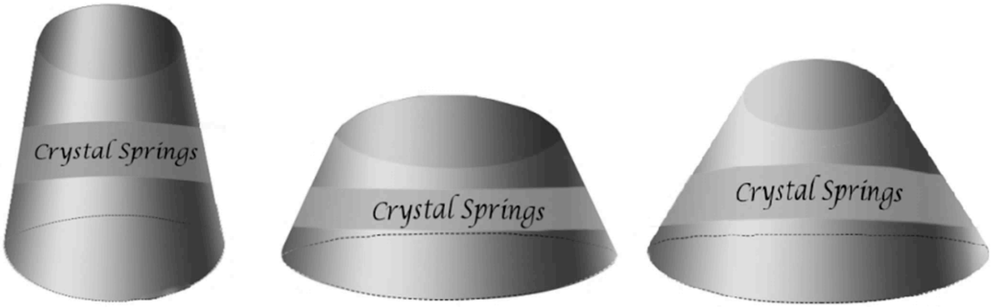

Figure 4. Sample stimuli from Experiment 3, varying in top-to-base ratio: 4:6 (left), 6:9 (middle), and 4:9 (right). Note the first two are low-disparity; the third is high-disparity.

Procedure. The procedure of Experiment 3 was identical to that of Experiment 2, except *all* participants were asked how much they could fill into these empty cups. It was emphasized that we were asking about the maximum amount the container could hold (i.e., that we meant filling to the brim). In addition, after providing each volume estimate, participants indicated the ease of performing the simulation by answering: “To what extent did you find it easy or difficult to mentally simulate filling up the cup?” (1 = Very Difficult; 9 = Very Easy).

Results

We defined four Level-1 variables: *orientation*, *shape, disparity*, and *height*. Orientation differentiated cups that were depicted right-side-up (+1) and upside-down (-1). Shape differentiated cups that were wide-based (+1) and wide-topped (-1). Disparity distinguished cavernous cups with the greatest disparity between their top and base (+2) from those with tops and bases that were more similar in size (-1). We also included height, a predictor that created an orthogonal contrast to disparity, but that distinguished between the three unique shapes of cup: the tallest cup (+1), the middle-height cup (0), and the shortest cup (-1).

All four Level-1 variables were nested within participant in a random-slope, random-intercept model. This permitted the effect of each factor to vary by participant (random-slope) and accounted for differences between participants in whether they tended to see all cups as generally larger or smaller (random-intercept). We included two two-way interactions: Shape X Disparity, given its importance to testing our explanation of the cavern hypothesis, and Shape X Height, given Height is the orthogonal contrast to Disparity. Finally, we included the random effect of *cup* used in each study. In the two sections below, we focus on different terms in the model to test the orientation and cavern effects:

Orientation effect. First, we investigated the orientation hypothesis. As in previous studies, we found that there was indeed a main effect of orientation on volume estimates, $t(5485.31) = 4.02$, $p < .001$. Replicating Experiments 1 and 2, we found that participants thought the same cup appeared to be larger when it was right-side-up (*M* = 16.51 xids) as opposed to upside-down (*M* = 16.03 xids).

To extend beyond this replication, we examined whether the ease of simulation explained the orientation effect. When predicting simulation ease with the same model, we found a main effect of orientation, *t*(158.20) = 3.99, *p* < .001. As expected, participants reported it was simpler to imagine filling the right-side-up cup (*M* = 5.63) than the upside-down one (*M* = 5.47). When we included simulation ease in our original model, we found that ease of simulation was positively related to volume estimates, *t*(115.82) = 4.46, *p* < .001, and the effect of orientation was reduced, but still significant, *t*(5377.61) = 3.69, *p* < .001. These results are consistent with simulation ease partially mediating orientation's effects on volume estimates, *z* = 2.97, *p* =.003. In short, the same cup looks lager right-side-up than upside-down in part because it is easier to simulate filling it up.

Cavern effect. We proceeded to investigate if the cavern effect—the tendency to judge the same container as larger when people see the base as large and the opening as small—depends on the disparity between the top and the base. Replicating Experiment 2, we found a main effect of shape that shows that we did observe a cavern effect overall, *t*(5203.78) = 2.99, *p* = .003. However, consistent with our central predictions, we observed a significant Shape X Disparity interaction, *t*(4374.89) = 4.28, *p* < .001. This demonstrated that the cavern effect emerged only when the top and base were sufficiently discrepant (4:9, 9:4), *t*(1544.04) = 4.82, *p* < .001. The cavern effect did not emerge when the discrepancy was smaller: regardless of whether we examined the short cups (6:9, 9:6), *t*(1434.39) = 1.75, *p* = .081, or the tall cups (4:6, 6:4), *t*(1563.79) = -1.40, *p* = .164. These effects, as well as a theoretically-irrelevant main effect of height (a proxy for shape), are depicted in Figure 5.

Our explanation of the cavern effect is not that it is easier to imagine filling up cups with low top-to-base ratios. Instead, we have emphasized that the contrast between the narrow opening that people imagine pouring into and the relatively vast base produces a cavernous illusion. Showing that ease of simulation did not also account for the cavern effect (as it did for the orientation effect), we found no evidence of the same Shape X Disparity interaction on ease of simulation that we found on volume estimation, $t < 1$.

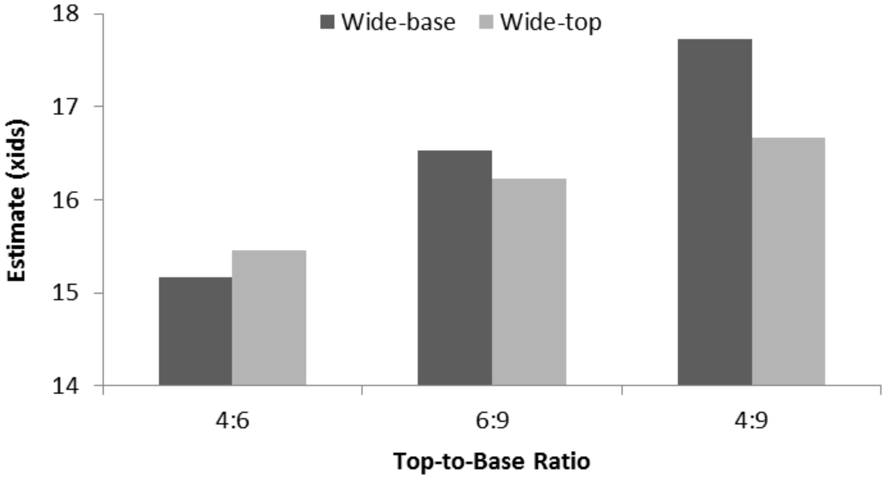

Figure 5. Mean volume estimates by base and height (indexed by top-to-base ratio) from Experiment 3. The cavern effect emerges only for cups with the greatest disparity between the top and base.

Experiment 4

Although Experiment 3 demonstrated that a relatively large discrepancy between a top and base was necessary for the cavern effect to emerge, the study confounded the shape of the cup with the ratio of the opening (through which one would fill the container) to the base. For example, the high-disparity cup—the only one that produced a cavern effect—was also the cup with the most shallowly-sloping sides (a consequence of the discrepancy between the top and the base). To address this limitation, Experiment 4 manipulated the size of the containers' top independently of the size of the opening. Some participants saw *open* cups—similar to those seen before—in which the opening was the entirety of the top. But other participants saw *lidded* cups, whose tops were covered except for a narrow hole that served as the opening. Lids narrowed the opening of all cups, regardless of their shape; thus, they should disrupt the cavern effect.

Method

Participants and design. Three hundred sixty-four participants took part in Experiment 4. Participants were randomly assigned to judge cups that were *open* (as before) or *lidded*. On 10 of the 8,307 trials (0.12%) for which participants offered a volume estimate, participants made an estimate of zero xids. As before, these trials were excluded.

Materials. We generated six sizes of cup (7.5 xids, 10 xids, 12.5 xids, 17.5 xids, 20 xids, and 22.5 xids). For wide-topped and wide-based cups, the top-to-base ratio was 4:9 and 9:4, respectively—the ratios that Experiment 3 found were sufficiently discrepant to produce the cavern effect. Half of participants saw open cups without lids (represented by open, black tops). The other half of participants saw the same cups that had lids with a small opening in the center, marked by a small black circle.

Procedure. Participants completed 12 trials in a randomized order, each consisting of viewing the modulus cylinder of 15 xids before estimating the volume of the target cup in xids.

Results

We defined one Level-1 variable: *shape*. As in Experiment 1, shape differentiated cups that were wide-based (+1) and wide-topped (-1). We defined a Level-2 variable, *opening*, which distinguished participants who saw cups with open tops (+1) and lidded tops (-1). The Level-1 variable was nested within participant in a random-slope, random-intercept model. This permitted the effect of each factor to vary by participant (random-slope) and accounted for differences between participants in the extent to which they saw all cups as generally bigger or smaller (random-intercept). We included the two-way interaction of shape and opening that was crucial for testing our primary hypothesis. Finally, we included the random effect of *cup.* (Note that in this study, this is merely a proxy for size.)

Supporting our central hypothesis, we observed a significant Base X Opening interaction, *t*(7249.49) = 2.85, *p* = .004 (see Figure 6). For open cups, we replicated the cavern effect, *t*(786.98) = 2.66, *p* = .008. But on lidded trials, when participants imagined pouring through a tiny opening for all cups, the cavern effect disappeared, *t*(786.81) = -0.65, *p* = .513 . Note that the absence of a cavern effect in the lidded cups was due to the narrow opening increasing the perceived volume of the wide-topped cups. Only our SIP account, not one based merely on shape, could predict this specific outcome.

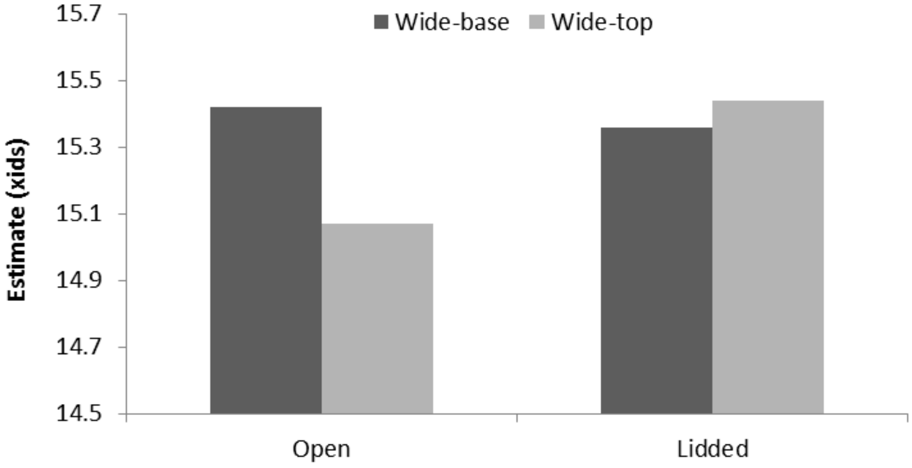

Figure 6. Mean volume estimates from Experiment 4. The absence of a cavern effect for lidded cups suggests it is the opening-to-base ratio that produces the cavern effect.

General Discussion

Four experiments support the hypothesis that simulation informs perception: People will estimate a container's volume by simulating how much they can fill into it. This approach helped identify two non-normative influences on volume estimation. First, a right-side-up cup appeared larger than the same one turned upside-down (Experiments 1-3). Second, the same container appeared larger when its narrow end was said to be its top, and its wide end, its base (Experiments 2-4). The emergence of both effects depended on the type of simulation conducted, and hence did not result from a different perception of the targets' low-level features (Experiment 2). The orientation effect resulted from the greater ease of simulating filling the right-side-up (vs. upside-down) cup (Experiment 3). The cavern effect emerged due to the sufficiently large disparity between a narrow opening (through which filling happens) and the relatively wide base (Experiments 3-4).

The current approach represents a significant departure not merely from most research on volume estimation, but also research on simulated judgment. Given that we identify a feature of the judgment target, not an ephemeral state of the perceiver (e.g., thirst; Balcetis & Dunning, 2006), that affects simulation, this may make the present findings easier to anticipate in natural settings. Differentiating our approach from past research on errors in volume estimation, we did not identify invalid algorithms that lead size evaluations astray, but instead tested how a cognitive procedure that underlies perceptual judgments produces predictable distortions.

Firestone and Scholl (F&S; in press) outline a checklist (italicized below) for future researchers who wish to claim that cognition does indeed distort perception. On the one hand, given our claim—i.e., that cognition is often the bridge between modular low-level perception and more complex perceptions—is less radical, our challenge is not necessarily to satisfy these six criteria. At the same time, it is worth considering how the present findings successfully circumvent or relate to those pitfalls. For example, Study 2 sidesteps two concerns. By holding the stimuli constant (thereby *not changing bottom-up features*) and varying the relevant simulation (emptying or filling), we *make disconfirmatory predictions* (of when the cavern and orientation effects should be attenuated). Our paradigms do not confuse perceptual judgments with those of *memory and recognition*. (Even if one wanted to propose that the greater size of right-side-up cups stems from their greater familiarity, one would have difficulty explaining why the more typical wide-topped cups look smaller than their wide-based counterparts.) Furthermore, Experiment A (Supplemental Materials) ruled out the influence of *demand effects.*

The final two concerns deserve more careful consideration. F&S note that sometimes cognition doesn't change perception, but instead calls that perception into question, leading to a reasoned shift in subsequent judgment. For example, throwing a heavy (vs. light) ball at a target increases estimates of the target's distance (Witt et al., 2004). Woods et al. (2009) offer additional evidence suggesting participants still saw the target as the same distance away, but merely disbelieved the visual input ("It looks close by, but it sure seemed hard to throw this over there…") In the present studies, however, participants were not given a new fact about their environments that caused them to reinterpret their initial perceptions. Instead, simulation was leaned upon spontaneously to inform volume perceptions.

Finally, F&S argue that some top-down influences on perception are more appropriately characterized as effects of attention. In an extreme case, people may literally shut their eyes, thereby altering what they perceive. We agree with F&S that "this familiar 'top-down' effect clearly isn't revolutionary…[it] changes only the *input* to perception, without changing perceptual processing itself" (p. 9). Nevertheless, the idea that attention may change *how* low-level perceptual features are ultimately integrated into a more complex percept also serves to reinforce our basic point that there is latitude in how perceptual input is integrated into more complex perceptual judgments. F&S say that “easily, the most natural and robust distinction between types of mental processes is that between perception and cognition” (p. 4). The present paper serves as one model for how these two distinct systems can still be intertwined, with perceptual input informing cognitive procedures that guide more complex perceptions.

References

Anderson, N. H., & Cuneo, D. O. (1978). The height+ width rule in children's judgments of quantity. *Journal of Experimental Psychology: General*, *107*, 335-378.

Balcetis, E., & Dunning, D. (2006). See what you want to see: motivational influences on visual perception. *Journal of Personality and Social Psychology*, *91*, 612.

Durgin, F. H., Klein, B., Spiegel, A., Strawser, C. J., & Williams, M. (2012). The social psychology of perception experiments: Hills, backpacks, glucose and the problem of generalizability. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 1582–1595.

Eibach, R. P., Libby, L. K., & Gilovich, T. D. (2003). When change in the self is mistaken for change in the world. *Journal of Personality and Social Psychology*, *84*, 917-931.

Firestone, C. (2013). How "paternalistic" is spatial perception? Why wearing a heavy backpack doesn't—and couldn't—make hills look steeper. *Perspectives on Psychological Science*, *8*, 455-473.

Firestone, C., & Scholl, B. J. (2014). "Top-down" effects where none should be found the El Greco fallacy in perception research. *Psychological Science*, *25*, 38-46.

Firestone, C., & Scholl, B. J. (2015). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Brain and Behavior Sciences*.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.

Gilbert, D. T., Gill, M. J., & Wilson, T. D. (2002). The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes*, *88*, 430-444.

Li, Y., Johnson, E. J., & Zaval, L. (2011). Local warming: daily variation in temperature affects beliefs and concern about climate change. *Psychological Science, 22,* 454-459.

Linkenauger, S. A., Witt, J. K., Stefanucci, J. K., Bakdash, J. Z., & Proffitt, D. R. (2009). The effects of handedness and reachability on perceived distance. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1649.

Ordabayeva, N., & Chandon, P. (2013). Predicting and managing consumers' package size impressions. *Journal of Marketing*, *77*, 123-137.

Proffitt, D. R. (2006). Distance perception. *Current Directions in Psychological Science*, *15*, 131-135.

Proffitt, D. R., Stefanucci, J., Banton, T., & Epstein, W. (2003). The role of effort in perceiving distance. *Psychological Science*, *14*, 106-112.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*, 341-365.

Raghubir, P., & Krishna, A. (1999). Vital dimensions in volume perception: Can the eye fool the stomach? *Journal of Marketing Research*, *36*, 313-326.

Risen, J. L., & Critcher, C. R. (2011). Visceral fit: While in a visceral state, associated states of the world seem more likely. *Journal of Personality and Social Psychology*, *100*, 777-793.

Schnall, S., Zadra, J. R., & Proffitt, D. R. (2010). Direct evidence for the economy of action: Glucose and the perception of geographical slant. *Perception*, *39*, 464-482.

Simmons, J. P., Nelson, L. D., & Simonsohn, Uri (2013). Life after P-Hacking (2013). Meeting of the Society for Personality and Social Psychology, New Orleans, LA. Available at SSRN: https://ssrn.com/abstract=2205186.

Stefanucci, J. K., Proffitt, D. R., Clore, G. L., & Parekh, N. (2008). Skating down a steeper slope: Fear influences the perception of geographical slant. *Perception*, *37*, 321-323.

Wansink, B., & Van Ittersum, K. (2003). Bottoms up! The influence of elongation on pouring and consumption volume. *Journal of Consumer Research*, *30*, 455-463.

Witt, J. K., & Proffitt, D. R. (2008). Action-specific influences on distance perception: a role for motor simulation. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1479.

Yang, S., & Raghubir, P. (2005). Can bottles speak volumes? The effect of package shape on how much to buy. *Journal of Retailing*, *81*, 269-281.

Supplemental Materials

Experiment A

We had three goals with Experiment A. First, we wanted to replicate both the orientation and cavern effects in a paradigm in which we did not ask participants to engage in a simulation. Although Experiment 1 was characterized by this feature, it tested the orientation effect only. Second, although Experiment 2 confirmed the key role of simulation in producing the orientation and cavern effects, we wanted to more carefully probe the role of attention in producing these effects. By our logic, the orientation affect does not relate to attention, but instead to ease of simulation. But people simulate filling a right-side-up and an upside-down container by scanning the image from top-to-bottom and from bottom-to-top, respectively. If it is the attentional pattern produces the effect, then having people process or scan containers from (visual) top-to-bottom (vs. bottom-to-top) should produce a similar inflation in people's estimates. In contrast, we propose the cavern effect *is* related to attention. That is, we argue that because people estimate a container's volume by simulating filling it from top to base, the stark contrast between a narrow top and a larger base produces a cavernous illusion. Though if people were exogenously prompted to process an image not as they naturally would (from its aperture to its base), but instead fro its narrow end to wide end (regardless of which is the top or the base), this might produce the same cavernous illusion.

Third, we wanted to determine whether our results could have been driven by demand effects. Toward this end, all subjects were asked to report how they thought the experimenters may have been using orientation and shape to bias volume estimates. We tested: 1) whether participants could intuit the orientation and cavern effects, and 2) whether accounting for such beliefs eliminated the orientation and cavern effects.

To test these ideas, we added two new conditions in which we prompted participants to process the images from the top to the bottom of the image (*scan-down* condition), or from the bottom to the top of the image (*scan-up* condition). If (not as hypothesized) the orientation effect and (as hypothesized) the cavern effect result from the dynamic sequencing of where participants allocate attention to and process the target, these manipulations should have predictable consequences on volume estimates. Namely, if the orientation effect is a function of attention, we should find volume estimates to be greater in the scan-down condition than the scan-up condition. If the cavern effect is a function of attention, we should find volume estimates to be greater when participants scan from the smaller base to the larger base. We pre-registered our hypotheses and materials with the Open Science Framework here: https://osf.io/th4ek/?view_only=4033077e42c1490687d41ac4ad79ddeb.

Participants and design. Six hundred sixteen Americans were recruited from Amazon's Mechanical Turk. Participants were randomly assigned to one of three processing conditions: *control, scan-down,* or *scan-up.*

Materials. The same 24 cups used in Experiment 2's fill condition were the target stimuli in Experiment A. As before, these cups were empty, and varied on non-crucial dimensions (color and size) as well as dimensions central to our hypotheses (orientation and shape). After participants made all 24 volume estimations, they were given this question stem: “"While making the estimates, I assumed the experimenters were predicting that by having a cup be __________, that this would make me judge the cup to be smaller/bigger than it actually was..." Participants responded to 6 features, 4 of which were hypothesis-relevant, on slider scales anchored at -50 (*much smaller*), 0 (*when making judgments, didn't think about experimenters* *predicted either way*), or 50 (*much bigger*): right-side-up, upside-down, green, purple, wide topped and narrow based, wide based and narrow topped.

Stimuli. There were two main differences between the procedure used in Experiment 2 and Experiment A. First, whereas in Experiment 2 we varied the nature of the simulation in which we asked participants to engage, in Experiment A we merely asked, “How many xids can this cup hold?”

Second, we presented the stimuli in one of three ways. In the control condition, participants saw a static image of the stimuli just as before. In the scan-down condition, the image loaded dynamically from top to bottom. In the scan-up condition, the image loaded dynamically from bottom to top. It took one second for the image to load. Two seconds later, the dynamic loading was repeated. At that point, the complete image remained until a volume estimate was provided.

Results. We began by testing whether we replicated both the orientation and cavern effects in our control condition (which made no reference to filling). Second, we examined whether our new scan-up and scan-down conditions could provide evidence that simulation changes the dynamics of how the targets are attended to, which might underlie the orientation effect (counter to hypotheses) or the cavern effect (consistent with hypotheses). Third, we tested whether participants intuited our hypotheses and whether such intuitions could suggest a demand effect that would explain our results.

Do participants spontaneously employ simulation? We began by restricting our analyses to the control condition, in which participants were merely asked how much the cup could hold. We defined orientation (+1: right-side-up, -1: upside-down) and shape (+1: wide-based, 0: cylindrical; -1: wide-topped) as two Level-1 variables nested within participant in a random-slope, random-intercept model. As before, we include a random effect of *cup* to account for theoretically-irrelevant variance attributable to other features (color and size).

We replicated both the orientation and cavern effects. Participants saw cups as being larger when right-side-up (M = 17.69 xids) than when upside-down (M = 17.12 xids), t(4997.61) = 3.73, $p < .001$. We ran the same model on the 16 non-cylindrical cups to test the cavern effect. And indeed, participants saw the targets as being larger when wide-based (M = 18.74 xids) than when wide-topped (M = 17.21 xids), t(3614.59) = 8.24, $p < .001$. That a right-side-up (vs. upside-down) orientation increased the perceived volume by 3.3% and that a wide-topped (vs. wide-based) shape increased the volume 8.9%—only somewhat smaller than the effects observed in fillcondition in Experiment 2: orientation effect, 4.8%; cavern effect, 10.4%—gives us confidence that many (and likely most) people spontaneously used simulation to estimate volume.

Does simulation operate through attention? We built on our basic model by defining a new Level-2 variable *processing,* a categorical variable that identified whether participants were assigned to the control, scan-down, or scan-up condition. We added the effect of processing, the Processing X Orientation interaction and the Processing X Shape interaction to the model. Given our interest in the processing condition, we no longer restricted our analyses to those in the control condition. Both interactions were significant: Processing X Orientation, F(2, 15384.54) = 3.10, p = .045; Processing X Shape, F(2, 15327.61) = 14.19, $p < .001$. These interactions both reflected that the orientation and shape effects were reduced in the scan-up and scan-down conditions. This is consistent with the idea that the two new conditions changed the way participants processed the images for the purpose of volume estimation. But did the disrupt the effects in a way consistent with the idea that one or both of our key effects operate through the effect of simulation on attention?

Orientation effect. We proceeded to test whether simulation's influence on the orientation and cavern effects operated through attention. To assess whether the orientation effect was plausibly an attentional effect, we looked at the main effect of processing. It achieved only marginal significance, $F(2, 642.59) = 2.40$, $p = .092$. But a closer inspection showed that the direction of the effect was opposite to that which would be expected if the orientation effect emerged due to attending to the (visual) top of an image before the base of an image. That is, participants who processed images by scanning down (from top to bottom) actually estimated that the images were slightly *smaller* ($M = 16.86$ xids) than those who scanned up ($M = 17.45$ *xids*), $t(646.64) = -1.97$, $p = .049$. Because the orientation effect was shown in Experiment 2 to be traceable to the metacognitive ease of completing the simulation, it should not be surprising that the orientation effect does not emerge from processing the top before the bottom of the visual field.

Cavern effect. But our account does implicate attention in the cavern effect. We defined a new variable: *scanned shape*. In the control condition, this variable took the same form as shape. But in the scan-up and scan-down condition, we assigned the value of +1 to trials on which participants were prompted to process the narrow end before the wide end and -1 to trails for which the wide base was processed before the narrow base.

For this analysis, we focused on the 16 trials for which participants judged a non-cylindrical shape. We tested a similar model to that which opened this section, but substituted Scanned Shape for Shape. In this case, the Processing X Scanned Shape interaction was significant, $F(2, 9251.71) = 6.42$, $p = .002$. The significant interacton reflected that although the cavern effect was significantly larger in the control condition, we were able to reproduce the cavern effect by prompting participants to attend to the narrow end before the wide end (in the scan-up and scan-down conditions). We conducted an additional model by combining across the two scanning conditions and using three Level-1 predictors: orientation, shape, and scanned shape. We observed a significant effect of shape, B = 0.17, *SE* = 0.06, *t*(6821.36) = 2.63, *p* = .009, reflecting that wide-based cups were judged larger than wide-topped cups even in the scan conditions. However, we observed a larger effect of scanned shape, B = 0.38, *SE* = 0.07, *t*(7096.91) = 5.84, *p* < .001. In other words, the dynamic stimulus loadings in the two scan conditions produced something akin to a cavern effect: Prompting participants to attend to the narrow end before the wide end (instead of the reverse) produced a cavern effect. This is consistent with the idea that attending to the narrow aperture before the wide base (the sequence of attention that would occur when simulating filling up a container) likely explains the cavern effect.

Do the orientation and cavern effects reflect demand effects? We first looked to participants' reports of their beliefs about the study's hypotheses. They showed a slight tendency to guess that the experimenters thought that the right-side-up cups would be judged larger (*M* = 5.17, *SD* = 14.41) than upside-down cups (*M* = 1.04, *SD* = 17.29), paired *t*(557) = 3.84, *p* < .001. This suggests that demand could have underlain the orientation effect, meaning that the data require further scrutiny. In contrast, the experimenters were assumed to think that the wide-based cups would be judged *smaller* (*M* = 3.29, *SD* = 19.70) than the wide-topped cups (*M* = 10.90, *SD* = 19.59), paired *t*(557) = -5.79, *p* < .001. Given this is the opposite of the actually hypothesized (and observed) effect, it is unlikely that demand underlies the cavern effect. Notably, both effects remained significant when we restricted our analyses to those in the control condition—i.e., those who most clearly showed our orientation and cavern effects.

To further probe whether the orientation effect may have been produced by experimental demand, we looked specifically at those participants who were in the control condition—i.e., those who showed the orientation effect. We defined a new variable, *orientation belief*, which was a difference score indicating participants' beliefs that the experimenters were expecting right-side-up cups would be judged larger than up-side-down cups. (Note that a difference score of 0 reflects a belief that orientation would have no effect.) We wanted to test whether the orientation emerged even among participants who did not guess the hypothesis—i.e., among those for whom their orientation belief was 0. We predicted estimates from orientation, shape, orientation belief, and the Orientation Belief X Orientation interaction. We continued to observe a significant effect of orientation, B = 0.28, $SE = 0.08$, $t(4840.64) = 3.31$, $p = .001$. Furthermore, the Orientation Belief X Orientation interaction was not significant, $t < 1$. In summary, although participants had some intuition behind the orientation (but not the cavern) effect, such intuitions were not traceable to their showing the orientation effect. This makes it highly unlikely that the orientation effect emerged to demand.

Additional Measures

Experiments 1-4. Beyond the primary measures reported in the main text, there were three additional measures collected. One was asked in each study. At the end of the study, participants saw a set of shapes and were asked which they had seen before. Participants were generally accurate in this judgment: 95.1% (Experiment 1), 89.4% (Experiment 2), 91.9% (Experiment 3), 95.8% (Experiment 4). We take no position on whether this memory question was sufficiently easy that it merely reflects that participants were paying attention, or whether it indicates generally high engagement with this study.

Experiment 4. In each study, participants completed quite a few trials. We wondered whether participants were actively engaging in the simulation on every trial, or whether they did not persist with this. If participants indicated they did always complete the simulations, it would be unclear if they felt the need to tell us that they were following instructions even if they weren't. But also, if participants indicated they did not always do this, it would suggest that the benefit we received from using a multi-trial, within-subjects design may not have been fully realized.

Thus, at the end of Experiment 4, we asked participants, "During this survey, while you did these mental simulations, to what extent did you imagine filling up the cup with water?" Granted, this question is somewhat ambiguous. Does it mean did they imagine filling it up all the way? Does it mean did they consistently do the simulation? Or does it mean did they complete the simulations vividly? Despite these questions about what the measure is capturing, it is somewhat heartening that only 8.6% of participants indicated they did not do this at all. Also, we saw that those on MTurk indicated they did this to a greater extent ($M = 5.48$, $SD = 1.41$) than those who completed the study in the lab ($M = 3.66$, $SD = 1.77$). MTurkers though often suspect that their compensation depends on their ability to demonstrate that they follow instructions. Our lab participants likely did not suspect their responses to this question would influence their receiving course credit. Thus, we caution readers from inferring that MTurkers are more engaged participants.

We asked a second question to our MTurk sample only: "When you imagined filling up the cup with water, did you imagine pouring into the top as it was depicted (either with the entire top open or with only a small hole open)?” Responses were made on a 7-point scale anchored at 1(*not at all*) and 7(*very much*). On the one hand, it is good that only 4.2% of participants indicated that they completely failed to follow this instruction. Participants generally indicated good compliance (*M* = 5.52, *SD* = 1.67). But as reviewed earlier, we are not certain that this measure is a valid indicator of the degree of compliance.

One approach for the future might be to use measures of this variety not to get precise estimates of absolute level of engagement, but to see what features of the experimental context might encourage more or less engagement. At the same time, researchers should be attentive to whether such moderators are actually moderators of the measures’ validity (i.e., by encouraging or discouraging truthful responding) instead of moderators of engagement.

More information on data and materials can be found here: https://osf.io/f67yr/?view_only=68a5f0aff1c342049ba5934307c626df