Research Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2019, Vol. 30(4) 495–503

DOI: 10.1177/0956797619831617
www.psychologicalscience.org/PS
SAGE

The Additive-Area Heuristic: An Efficient but Illusory Means of Visual Area Approximation

Sami R. Yousif and Frank C. Keil
Department of Psychology, Yale University

Abstract

How do we determine how much of something is present? A large body of research has investigated the mechanisms and consequences of number estimation, yet surprisingly little work has investigated area estimation. Indeed, area is often treated as a pesky confound in the study of number. Here, we describe the *additive-area heuristic*, a means of rapidly estimating visual area that results in substantial distortions of perceived area in many contexts, visible even in simple demonstrations. We show that when we controlled for additive area, observers were unable to discriminate on the basis of true area, per se, and that these results could not be explained by other spatial dimensions. These findings reflect a powerful perceptual illusion in their own right but also have implications for other work, namely, that which relies on area controls to support claims about number estimation. We discuss several areas of research potentially affected by these findings.

Keywords

visual perception, numerical cognition, size discrimination, open data, open materials, preregistered

Received 5/26/18; Revision accepted 10/25/18

When you look at a basket of oranges, a glass of water, or a piece of cake, how do you know how much is there? A great deal of research has investigated the shared capacity of adults, infants, and nonhuman animals to estimate the number of objects in a set (Barth, Kanwisher, & Spelke, 2003; Brannon & Terrace, 1998; Gordon, 2004; Nieder & Miller, 2004; Pica, Lemer, Izard, & Dehaene, 2004). This propensity to estimate large numerosities without counting is said to rely on an evolutionarily ancient system: the approximate number system (Halberda, Mazzocco, & Feigenson, 2008). Yet to our evolutionary ancestors, estimates of number may not have been the best assessment of amount.

Imagine foraging for food. Would you decide to forage from the bush with twice as many berries or the one with berries 3 times in volume? In many natural settings, size estimation rather than number estimation might be most critical for survival, although only a few studies have investigated approximate area perception conjointly with approximate number perception (Brannon, Lutz, & Cordes, 2006; Lourenco, Bonny, Fernandez, & Rao, 2012; Odic, Libertus, Feigenson, & Halberda, 2013). In fact, most studies have discussed area only in an attempt to rule out continuous spatial dimensions (e.g., area, contour length, density) as explanations for approximate number estimation (Barth, 2008; Mix, Huttenlocher, & Levine, 2002). Yet area or size perception is also an autonomous area of study: Models of area perception have been proposed in the context of development (e.g., Anderson & Cuneo, 1978; Gigerenzer & Richter, 1990), ensemble perception (e.g., Marchant, Simons, & de Fockert, 2013; Solomon, Morgan, & Chubb, 2011), perception research more broadly (e.g., Carbon, 2016; Ekman & Junge, 1961; Nachmias, 2008, 2011; Teghtsoonian, 1965), and even consumer decision making (e.g., Krider, Raghubir, & Krishna, 2001). Here, similarly, we demonstrated that area estimation itself reveals powerful and counterintuitive effects. Yet unlike the authors of much of the prior work, we assessed area perception (a) in the

Corresponding Author:
Sami R. Yousif, Yale University, Department of Psychology, Box 208205, New Haven, CT 06520-8205
E-mail: sami.yousif@yale.edu

context of numerous objects and (b) using displays akin to those commonly used to assess approximate number perception (e.g., Halberda et al., 2008; Odic et al., 2013). We showed that even in such displays, area estimation employs a simple heuristic that results in substantial distortions of perceived area. These distortions are not only important to understand in their own right; they also raise questions about attempts to control for area in number-estimation tasks. In particular, controlling for true area, insofar as it is dissociable from perceived area, may amplify a confound with numerosity in many studies.

The Additive-Area Heuristic

We propose that visual area estimation in simple visual displays is best captured by a single, simple heuristic: the additive-area heuristic. Consider Figure 1a. In which panel does it look like the circles cover a greater area: the left or the right? Although it may appear that the circles in the left panel cover a greater area than those in the right panel, the cumulative area covered by the circles in the two panels is equal. However, the panels differ in one important way: *Additive area* (i.e., the sum of a shape's dimensions rather than the product) is greater for the image on the left.

Five experiments were designed with the aim of manipulating either true area or additive area while holding the other constant. We found that (a) humans use a simple heuristic to calculate area, (b) humans often fail to perceive true area when accounting for this heuristic, (c) this heuristic cannot be explained by appeal to other dimensions, and (d) differences between true area and perceived area may have serious consequences for studies that rely on area as a manipulation or a control.

All of the experiments were preregistered. In addition to preregistering the sample size, basic methodology, and analysis plans, we also preregistered some details about how the stimuli were created. Note, however, that some of the language in the preregistration is slightly different from the language in this article. The preregistrations are available on the Open Science Framework (OSF) at osf.io/dc5t8.

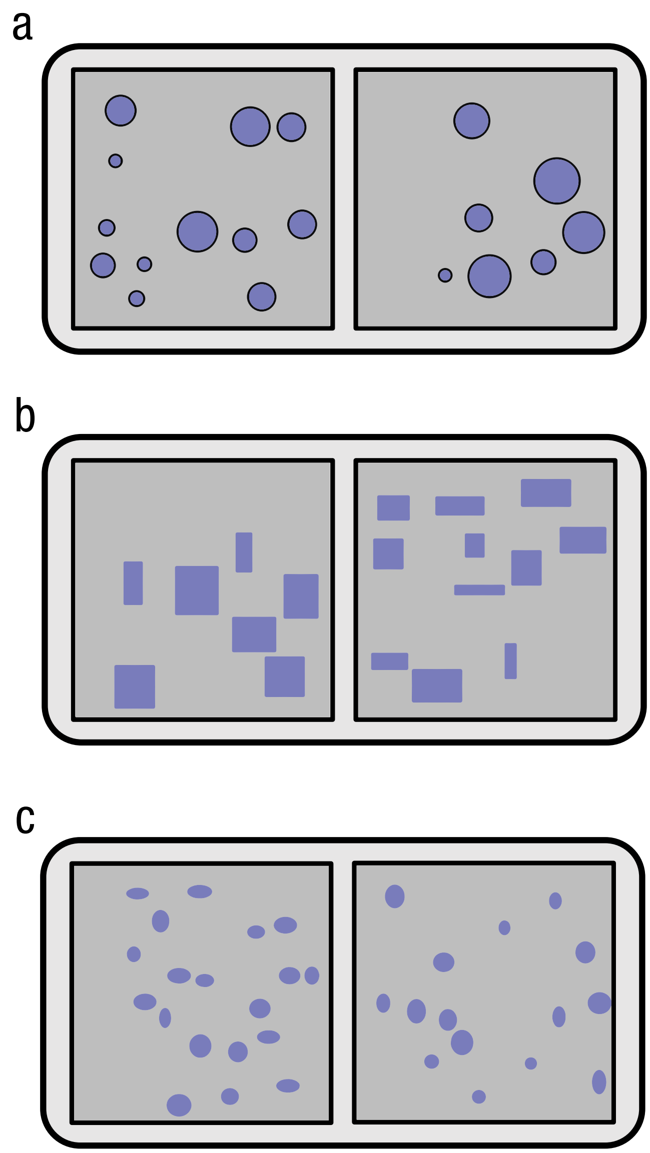

Fig. 1. Depiction of example displays from (a) Experiments 1, 2, and 4; (b) Experiment 3; and (c) Experiment 5. True area is equated for each pair in (a) and (b). However, additive area is 30% greater in the left panel of (a) and 30% greater in the right panel of (b). Perimeter is equated for each pair in (c). However, additive area is 30% greater in the left panel of (c). The stimuli appear here exactly as they would have to observers in the task. Additive area in each case is equal to the sum of the objects' height and width. For circles, additive area for each shape is equal to twice the diameter (which can be simplified to just diameter). For the rectangles, additive area for each shape is equal to height plus width. For ellipses, additive area for each shape is equal to height plus width (also the sums of the lengths of the major and minor axes).

Experiment 1: Additive Area Versus True Area

In the first test of the additive-area heuristic, observers completed an approximate-area task on displays of circular disks (see Fig. 1a). Critically, we varied these displays in terms of their cumulative true area as well as their cumulative additive area. We predicted that observers would be both slower to respond and less accurate when true area differed and that they would be faster to respond and more accurate when additive area differed.

Method

Observers. One hundred observers were recruited via Amazon's Mechanical Turk (MTurk), although 3 observers were excluded because they did not complete a single trial (i.e., they accepted but never started the task). All observers consented prior to participation, and the experiment was approved by the institutional review board at Yale University.

Materials. All of the stimuli were generated via custom software written in Python with the PsychoPy libraries (Peirce, 2007). The aim was to create pairs of stimuli that varied in either additive area or true area while the other was equated. For each stimulus pair, we randomly generated an initial set of disks (20–100 pixels in diameter, with a buffer of at least 10 pixels between any two disks) and then pseudorandomly generated a second set of objects on the basis of a given additive-area ratio. The initial set of objects always had seven disks. Stimulus pairs were generated randomly until a pair met both the additive-area criterion and the true-area criterion, at which point that pair would be rendered another time and saved. The second stimulus (i.e., not the set with seven disks) always had more area (whether additive area or true area) than the initial stimulus. Number was unconstrained in the stimulus-generation process, meaning that the number ratio was not equated across all possible additive-area and true-area ratios (1.5 on average for additive-area trials; 0.8 on average for true-area trials). For details on how additive area, true area, and number covaried, see the "Stimulus Details" files on the OSF (osf.io/dc5t8). All disks were rendered with a thin, black border (4-pixel stroke width). The images depicted in Figure 1 are representative, as they were actual images used in the experiments.

In this initial experiment, there were only two constraints: additive area and true area. There were pairs in which true area was equal (to serve as a baseline), pairs in which true area varied and additive area was controlled, and pairs in which additive area varied and true area was controlled. When additive area was controlled, true area could vary in a 1.00, 1.10, 1.20, or 1.30 ratio (and vice versa for additive area when true area was controlled). Because there are mathematical constraints on how much additive area and true area can differ, these ratios were selected to maximize the differences between them. Because of the pseudorandom nature of stimuli creation and the mathematical constraints involved in creating such stimuli, true area was never perfectly matched with the stated ratio; it could vary by 1% in either direction. That is, if the true-area ratio for a given trial were 1.10, we allowed the difference in true area to fluctuate between 1.09 and 1.11.

Procedure. The task was administered online via MTurk using custom software. On each trial, observers saw two spatially separated sets of lavender-colored dots, presented side-by-side in the center of the screen, with 50 pixels of space between each set. Each stimulus was 400 pixels × 400 pixels. The side that contained the set with more cumulative area (left vs. right) was counterbalanced. Observers were instructed to press "q" if the image on the left had more cumulative area and "p" if the image on the right had more cumulative area. Observers were told the following: "Your task is simply to indicate which set of circles has **more cumulative area**. In other words: if you printed the images out on a sheet of paper, which would require more total ink?" Later, they were told, "The sets of dots will sometimes vary in number, but the number of dots does not matter. Instead, you should answer only which has more area, regardless of number." The stimuli stayed on the screen until the observer responded, and there was no time limit on responses. Between each trial, there was a 1,000-ms intertrial interval. Observers completed 84 trials, 12 each of seven trial types (true area varying in a 1.10, 1.20, or 1.30 ratio while additive area was held constant; additive area varying in a 1.10, 1.20, or 1.30 ratio while true area was held constant; and both additive and true area being equal). All trials were presented in a unique random order for each observer. Observers completed two representative practice trials before beginning the task. Because more than half of the trials had no objectively correct answer (because true area did not vary), we measured accuracy as a propensity to choose "more," whether that be more additive area or more true area.

Results

The results are shown in Figure 2. Observers were indeed faster and more accurate in making discriminations on the basis of additive area rather than true area. A repeated measures analysis of variance (ANOVA) conducted on accuracy with two factors (condition: additive area, true area; ratio: 1.10, 1.20, 1.30) revealed main effects of both condition, $F(1, 96) = 17.80$, $p < .001$, and ratio, $F(2, 95) = 24.43$, $p < .001$, as well as an interaction between the two, $F(2, 95) = 9.94$, $p < .001$. Post hoc tests revealed that overall performance was above chance in the additive-area condition, $t(96) = 10.88$, $p < .001$, $d = 1.11$. However, surprisingly, observers were unable to make discriminations on the basis of true area alone, $t(96) = 1.70$, $p = .09$, $d = 0.17$. Even in the trials with the biggest difference in area (1.30 ratio), observers were not above chance in their area discriminations, $t(96) = 1.93$, $p = .06$, $d = 0.20$. A separate ANOVA conducted on response times revealed a similar pattern, and post hoc tests confirmed that observers were more than 120 ms

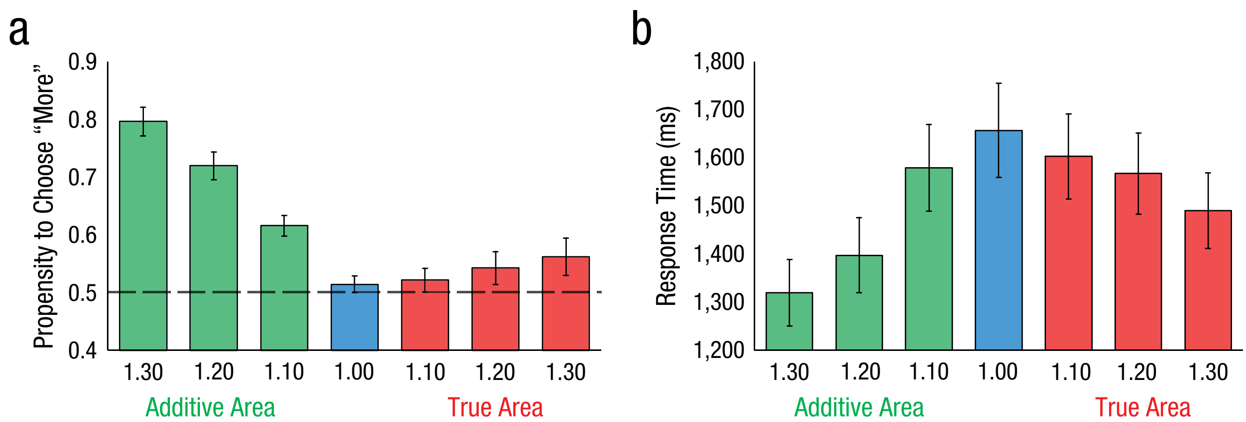

Fig. 2. Results from Experiment 1. The proportion of trials on which observers selected the option with "more"—whether that was more true area or more additive area—(a) is shown for each of the seven additive-area and true-area ratios tested. The dashed line represents chance performance. Mean response time (b) is shown for each of the seven ratios tested. In both graphs, the *x*-axis represents the ratio. When additive area varied, true area remained constant. When true area varied, additive area remained constant. Thus, green bars correspond to additive-area trials, red bars correspond to true-area trials, and the blue bar represents trials in which both areas were equal. Error bars represent ±1 *SE*.

faster for the additive-area trials compared with the true-area trials, $t(96) = 4.88$, $p < .001$, $d = 0.50$.

Because we did not explicitly manipulate number, we tested whether number could potentially explain these results. A linear regression with additive area, true area, and number as covariates revealed that additive area did predict observer responses ($p < .001$) but that neither true area ($p = .86$) nor number ($p = .23$) did. Results from a supplemental version of this experiment in which numerosity was further manipulated can be found on the OSF (Experiment S1; osf.io/dc5t8).

Discussion

Additive area can explain variance in area estimation, and surprisingly, observers seem unable to make discriminations using true area when additive area is controlled.

Experiment 2: Time-Limited Approximations

To ensure that all observers spent roughly the same amount of time assessing the displays and that these judgments were, in fact, rapid approximations, we replicated Experiment 1, except that observers had only 700 ms to view the stimuli.

Method

One hundred observers were recruited via MTurk, although 3 observers were excluded because they did not complete a single trial (i.e., they accepted but never started the task). All observers consented prior to participation, and the experiment was approved by the institutional review board at Yale University. The details of this experiment were identical to those of Experiment 1, except that the stimuli appeared for only 700 ms before disappearing. Observers still had unlimited time to respond.

Results

Results from this manipulation can be seen in Figure 3a. Observers were more accurate in the additive-area condition than the true-area condition, $t(96) = 6.40$, $p < .001$, $d = 0.65$, although there were no differences in response time (which is to be expected because observers were rushed), $t(96) = 1.60$, $p = .112$, $d = 0.50$. Observers were unable to make discriminations on the basis of true area alone, $t(96) = 1.26$, $p = .21$, $d = 0.13$.

Once again, linear regression revealed that additive area significantly predicted observer responses ($p < .005$), whereas number did not ($p = .51$). True area was also a significant predictor in this model ($p = .035$), although in the opposite direction (i.e., observers were less likely to choose the option with more area). This latter result is likely driven by all the ratios for which true area was zero and additive area varied and thus should not be overinterpreted. Note that overall performance for the true-area trials did not significantly differ from chance ($p = .21$).

Discussion

Experiment 2 further supports the heuristic nature of this phenomenon: In addition to replicating the results of Experiment 1, these results show that additive area was used for rapid approximation of visual displays.

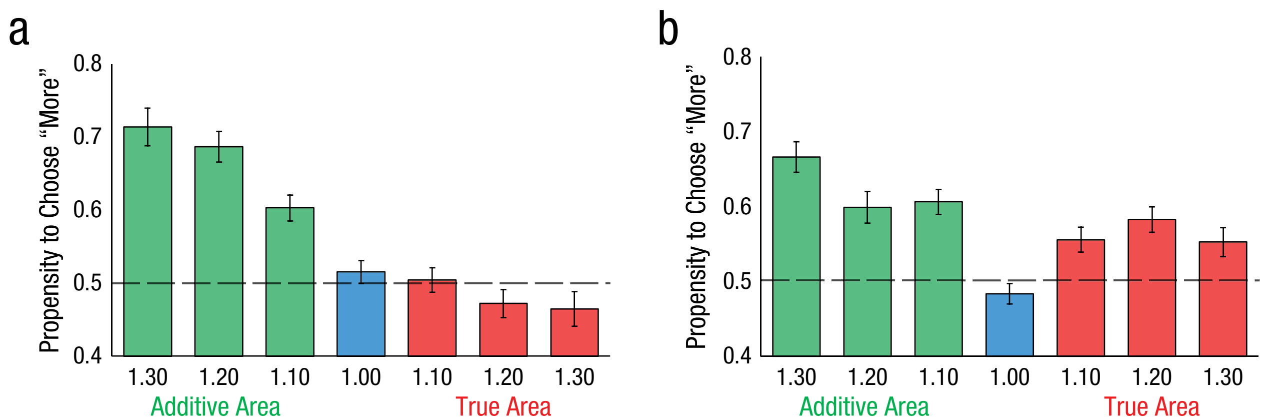

Fig. 3. Results from (a) Experiment 2 and (b) Experiment 3. The proportion of trials for which observers selected the option with "more"—whether that was more true area or more additive area—is shown for each of the seven additive-area and true-area ratios tested. The *x*-axes represent the ratio. When additive area varied, true area remained constant. When true area varied, additive area remained constant. Thus, green bars correspond to additive-area trials, red bars correspond to true-area trials, and the blue bar represents trials in which both areas were equal. The dashed lines represent chance performance. Error bars represent ±1 *SE*.

Experiment 3: Rectangles

Most studies on the approximate number system have relied on displays of disks. However, to ensure that the additive-area heuristic was not specific to such displays, we replicated the results of the prior experiment with rectangles instead of disks (see Fig. 1b).

Method

One hundred observers were recruited via MTurk, although 1 observer was excluded for not completing a single trial (i.e., they accepted but never started the task). All observers consented prior to participation, and the experiment was approved by the institutional review board at Yale University. The details of this experiment were identical to those of Experiment 2, except that the stimuli were rectangles instead of disks. The aspect ratio of the rectangles varied from 1.0 to 5.0 (minimum length = 20 pixels; maximum length = 100 pixels). (And to rule out minor differences caused by how the borders are rendered, we rendered these rectangles without borders.)

Results

Once again (see Fig. 3b), observers were more accurate in making discriminations on the basis of additive area rather than true area, $t(98) = 2.61$, $p = .01$, $d = 0.26$. Observers' true-area discriminations were above chance, $t(98) = 4.88$, $p < .001$, $d = 0.49$, although they made the correct selection only 57% of the time. A linear regression revealed that whereas both additive area ($p < .005$) and true area ($p < .05$) significantly predicted observer responses, number did not ($p = .61$).

Discussion

In general, observers had more trouble making area discriminations with rectangles both in additive-area and in true-area trials. We suspect that this is because one dimension was overweighted relative to the other, meaning that a slightly more complex model might best explain area approximations in such cases. Although the effects in this experiment were weaker than those in prior experiments, additive area still outperformed true area as a model of area approximation.

Experiment 4: Number Control

Might these results be explained by a confound with number? The creation of the stimuli in Experiments 1 to 3 was constrained in such a way that number was partially confounded with additive area. In all three cases, differences in additive area predicted accuracy, whereas differences in number did not. In a stronger test, we constructed stimuli for which we could independently manipulate number.

Method

One hundred observers were recruited via MTurk, although 2 observers were excluded because they did not complete a single trial (i.e., they accepted but never started the task). All observers consented prior to participation, and the experiment was approved by the institutional review board at Yale University. The details of this experiment were identical to those of Experiment 1, except as otherwise noted.

There were 60 pairs of stimuli, 4 each of 15 types (5 area ratios × 3 numerosities). Additionally, it should be

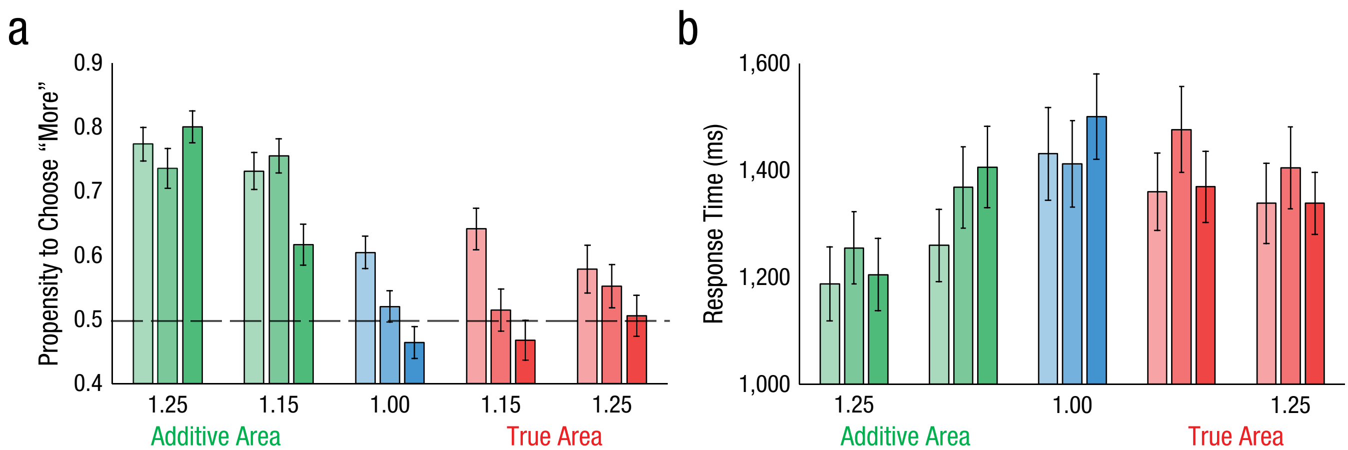

Fig. 4. Results from Experiment 4. The proportion of trials for which observers selected the option with "more"—whether that was more true area or more additive area—(a) is shown for each of the seven additive-area and true-area ratios tested. The dashed line represents chance performance. Mean response time (b) is shown for each of the seven ratios tested. In both graphs, the *x*-axis represents the ratio. When additive area varied, true area remained constant. When true area varied, additive area remained constant. Thus, green bars correspond to additive-area trials, red bars correspond to true-area trials, and the blue bar represents trials in which both areas were equal. Three different numerosities were tested for each area ratio; lighter bars correspond to lower numerosities within each set. Error bars represent ±1 *SE*.

noted that it was not possible to use the exact same number ratios across the additive area and true area. Instead, the goal was merely to have three different levels of numerosity at each additive-area and true-area ratio. This allowed us to independently assess the role of number at each level. To determine what numbers ought to be chosen in the first place, we ran an initial simulation to see how number would naturally vary (if unconstrained) for each additive-area and true-area ratio. From these initial simulations, we picked three of the possible numerosities. We purposefully chose numerosities that would maximally overlap across conditions (to minimize the impact of any unforeseen confound). The default number of items in each display was set to 10. Full stimulus details (including numerosities for every stimulus used in the experiment) can be found on the OSF (osf.io/dc5t8).

Results

The results from this experiment can be seen in Figure 4. We found that observers were both faster and more accurate in the additive-area trials compared with the true-area trials—accuracy: $t(97) = 5.02$, $p < .001$, $d = 0.51$; response time: $t(97) = 3.31$, $p = .001$, $d = 0.33$, replicating our previous results. Regression once again revealed that additive area ($p < .001$) but not true area ($p = .14$) significantly predicted observers' responses. The same regression revealed that number did significantly predict responses ($p = .004$); thus, greater numerosity resulted in a decreased likelihood of an observer indicating that an item had more area. However, this effect was specific to the true-area trials ($p = .004$). When we analyzed only the trials in which additive area varied, there was no effect of number ($p = .92$). In other words, although number may be used as a cue in certain contexts, it has no apparent effect on area judgments when perceived area differs.

Discussion

These results once again reveal the use of an additive-area heuristic. However, there was an effect of numerosity whereby the presence of additional disks in the display decreased the likelihood that an observer would indicate that the display had more area. This is in contrast to previous results that suggested correspondences between number and continuous magnitudes such as area (Hurewitz, Gelman, & Schnitzer, 2006). Thus, it seems that many past studies reporting influences of numerosity in these sorts of tasks may have been detecting variation caused by additive area instead. Importantly, when perceived area varied, observers did not rely on number as a cue.

Experiment 5: Perimeter Control

More than most continuous spatial dimensions, perimeter has been a dimension of interest in these sorts of displays (e.g., DeWind, Adams, Platt, & Brannon, 2015; McCrink & Wynn, 2007), and there is some evidence that perimeter may actually explain number approximation (see DeWind et al., 2015; Mix et al., 2002). Although perimeter-based approximations may not serve as feasible models of area perception (see the coastline paradox; Mandelbrot, 1967), they should be addressed. In the final experiment, we used a new stimulus to fully dissociate perimeter from perceived area: ellipses.

Method

One hundred observers were recruited via MTurk. All observers consented prior to participation, and the experiment was approved by the institutional review board at Yale University. The details of this experiment were identical to those of Experiment 2, except that we substituted perimeter for true area. We used ellipses in place of disks. In other words, additive area varied while perimeter was held constant, and perimeter varied while additive area was held constant (and both varied in 1.0, 1.1, 1.2, and 1.3 ratios). No specific limits were imposed on area or numerosity (meaning that, in practice, they varied much more than either additive area or perimeter). The default number of stimuli was 15. These stimuli were rendered without borders. The aspect ratio of the disks ranged from 1.00 to 2.20.

Results

Overall, observers were better at making discriminations on the basis of additive area than on the basis of perimeter, $t(98) = 10.31$, $p < .001$, $d = 1.04$. Observers were unable to make discriminations on the basis of perimeter alone, $t(98) = 0.99$, $p = .33$, $d = 0.10$.

Discussion

Cumulative perimeter was unable to explain the approximation of area, whereas additive area alone was able to do so.

General Discussion

Not only does the additive-area heuristic account for a high proportion of the variance in area judgments, but observers also seem to be insensitive to differences in true area under certain conditions. These results have implications for many different research programs in cognitive science. We highlight four areas of active research likely to be influenced by these findings.

Visual perception

Many researchers have addressed the question of size perception (Ekman & Junge, 1961; Teghtsoonian, 1965). Some have addressed illusions of visual size (Coren & Girgus, 1978). Others have discussed the continuous dimensions of space that influence the perception of not only size but also density, numerosity, and texture (e.g., Anobile, Cicchini, & Burr, 2014, 2016; Durgin, 1995). In all of these cases, the additive-area heuristic offers a simple, powerful, low-dimensional means of area estimation. This finding may clarify and unify various prior studies on the perception of area (e.g., Carbon, 2016) while also raising questions about links between the perception of size (of a single object), area (of a set of objects), density, and texture.

Approximate area

The study of approximate area is not nearly as pervasive as the study of approximate number, yet the authors of several prominent articles have studied the two in tandem (Lourenco et al., 2012; Odic et al., 2013). Both approximate number estimation and approximate area estimation are proposed to independently contribute unique variance to mathematical competence (Lourenco et al., 2012). However, this work involved manipulation of mathematical area rather than perceived area. Thus, number discrimination could have been influenced by additive area, even though mathematical area was controlled.

Approximate number

The subject of hundreds of articles and cumulatively tens of thousands of citations, the approximate number system has dominated the field of numerical cognition for the past decade (Barth et al., 2003; Halberda et al., 2008; Lourenco & Bonny, 2017; Lourenco et al., 2012). Much attention has been given to the continuous spatial dimensions that are confounded with numerosity (e.g., Barth, 2008; DeWind et al., 2015; Mix et al., 2002). Of these, area is by far the most common control (e.g., Halberda et al., 2008; Lourenco et al., 2012; Xu & Spelke, 2000). Yet if true area is different from perceived area, variance in perceived area might well explain performance on these tasks.

General magnitude

Several researchers have investigated the link between number and other magnitudes. One prominent theory suggests that representations of time, space, quantity, and other magnitudes rely on similar cortical processes (Lourenco & Longo, 2010; Sokolowski, Fias, Ononye, & Ansari, 2017; Walsh, 2003). In support of this theory, many researchers have pointed to Stroop-like errors between area and number (Brannon, Abbott, & Lutz, 2004; Hurewitz et al., 2006; Rousselle, Palmers, & Noël, 2004). Although the present results do not bear on all facets of this extensive literature, they do relate to the tendency to use number as a cue to approximate area and vice versa (e.g., Hurewitz et al., 2006). This could be the result of shared mechanisms, but it might also be the result of a simple confound. The bias to select the set with greater numerosity might instead be a bias to select the set with more perceived area. Our findings suggest that either (a) number has an adverse effect on area estimation—exactly the opposite of the general-magnitude account—or (b) observers are using some other heuristic to make their responses in these cases (e.g., choosing the display with the single largest object).

The Illusion of Approximate Area

There have been many careful attempts to capture approximate number acuity by modeling a nearly exhaustive list of continuous dimensions of the stimuli (DeWind et al., 2015), with researchers concluding that continuous dimensions of space influence the approximation of number. What does our approach reveal that is not already captured by existing models? Consider the Ebbinghaus illusion, whereby one disk, surrounded by many smaller disks, appears greater in size than an equal-size disk surrounded by many larger disks. Modeling approximate-number-system performance by exhaustively characterizing every continuous dimension of a display is akin to explaining the Ebbinghaus illusion by measuring every continuous dimension of the two disks being compared. No measurements collected on the relevant disks could explain the Ebbinghaus illusion because they are exactly the same; it can be explained only by appeal to perception.

By contrast, we made an explicit prediction about what drives the area approximation and manipulated that specific dimension to eliminate differences in perceived area. This does not mean that additive area fully explains area perception: There may be context- or task-dependent interactions among many continuous variables (e.g., density, convex hull, average element size) that contribute to the perception of size. Yet controlling the additive area eliminated the ability to distinguish displays in most cases, providing strong evidence that this factor is directly linked to area perception. Further, this heuristic offers a simple solution that may be easily implemented in studies on the approximate number system.

Conclusion

This article documented the additive-area heuristic, a simple, low-dimensional heuristic that accounts for substantial variability in area approximation. The explanatory power of this heuristic persists despite variance in other salient dimensions (e.g., true area, perimeter, and number) and may bear on the interpretation of many seminal articles in the field of numerical cognition as well as work on area estimation, general magnitude, and various aspects of visual perception. The notion of perceived area helps explain other findings in many diverse fields of cognitive science while advancing those fields both theoretically and methodologically.

Action Editor

Alice J. O'Toole served as action editor for this article.

Author Contributions

S. R. Yousif developed the study concept and the study design with input from F. C. Keil. Testing, data collection, and data analysis were performed by S. R. Yousif with input from F. C. Keil. S. R. Yousif drafted the manuscript, and F. C. Keil provided critical revisions. Both authors approved the final manuscript for submission.

Acknowledgments

We thank Brian Scholl, Richard Aslin, Justin Halberda, Brynn Sherman, April Bailey, Stefan Uddenberg, Ben van Buren, Emma Alexandrov, Elizabeth Bennette, and all the members of the Yale Perception and Cognition Lab and the Yale Cognition and Development Lab for helpful conversations. We thank Darko Odic, Richard Aslin, Nora Newcombe, and Lauren Aulet for comments on earlier drafts of the manuscript.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This project was supported by a National Science Foundation Graduate Research Fellowship awarded to S. R. Yousif.

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at osf.io/dc5t8. The design and analysis plans for the experiments were preregistered (osf.io/dc5t8). The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619831617. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

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.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science*, *25*, 265–270.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number as a primary perceptual attribute: A review. *Perception*, *45*, 5–31.

Barth, H., Kanwisher, N., & Spelke, E. S. (2003). The construction of large number representations in adults. *Cognition*, *86*, 201–221.

Barth, H. C. (2008). Judgments of discrete and continuous quantity: An illusory Stroop effect. *Cognition*, *109*, 251–266.

Brannon, E. M., Abbott, S., & Lutz, D. J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition*, *93*, 59–68.

Brannon, E. M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental Science*, *9*, 59–64.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science*, *282*, 746–749.

Carbon, C.-C. (2016). The folded paper size illusion: Evidence of inability to perceptually integrate more than one geometrical dimension. *i-Perception*, *7*(4). doi:10.1177/2041669516658048

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusions*. Hillsdale, NJ: Erlbaum.

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265.

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 149–169.

Ekman, G., & Junge, K. (1961). Psychophysical relations in visual perception of length, area and volume. *Scandinavian Journal of Psychology*, *2*, 1–10.

Gigerenzer, G., & Richter, H. R. (1990). Context effects and their interaction with development: Area judgments. *Cognitive Development*, *5*, 235–264.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*, 496–499.

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668.

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences, USA*, *103*, 19599–19604.

Krider, R. E., Raghubir, P., & Krishna, A. (2001). Pizzas: π or square? Psychophysical biases in area comparisons. *Marketing Science*, *20*, 405–425.

Lourenco, S. F., & Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Developmental Science*, *20*, 12418–12420.

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences, USA*, *109*, 18737–18742.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*, *21*, 873–881.

Mandelbrot, B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, *156*, 636–638.

Marchant, A. P., Simons, D. J., & de Fockert, J. W. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta Psychologica*, *142*, 245–250.

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, *18*, 740–745.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin*, *128*, 278–294.

Nachmias, J. (2008). Judging spatial properties of simple figures. *Vision Research*, *48*, 1290–1296.

Nachmias, J. (2011). Shape and size discrimination compared. *Vision Research*, *51*, 400–407.

Nieder, A., & Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences, USA*, *101*, 7457–7462.

Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology*, *49*, 1103–1112.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*, 499–503.

Rousselle, L., Palmers, E., & Noël, M. P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology*, *87*, 57–84.

Sokolowski, H. M., Fias, W., Ononye, C. B., & Ansari, D. (2017). Are numbers grounded in a general magnitude processing system? A functional neuroimaging meta-analysis. *Neuropsychologia*, *105*, 50–69.

Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *Journal of Vision*, *11*(12), Article 13. doi:10.1167/11.12.13

Teghtsoonian, M. (1965). The judgment of size. *The American Journal of Psychology*, *78*, 392–402.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.