Received: 19 July 2021 | Revised: 22 November 2021 | Accepted: 6 January 2022

DOI: 10.1111/desc.13235

Developmental Science WILEY

RESEARCH ARTICLE

Do children estimate area using an "Additive-Area Heuristic"?

Sami R. Yousif[1] | Emma Alexandrov[1] | Elizabeth Bennette[1] | Richard N. Aslin[1,2,3] | Frank C. Keil[1]

[1] Department of Psychology, Yale University, New Haven, Connecticut, USA

[2] Haskins Laboratories, New Haven, Connecticut, USA

[3] Yale Child Study Center, New Haven, Connecticut, USA

Correspondence
Sami Yousif, Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, USA.
Email: sami.yousif@yale.edu

Abstract

A large and growing body of work has documented robust illusions of area perception in adults. To date, however, there has been surprisingly little in-depth investigation into children's area perception, despite the importance of this topic to the study of quantity perception more broadly (and to the many studies that have been devoted to studying children's number perception). Here, in order to understand the interactions of number and area on quantity perception, we study both dimensions in tandem. This work is inspired by recent studies showing that human adults estimate area via an "Additive Area Heuristic," whereby the horizontal and vertical dimensions are summed rather than multiplied. First, we test whether children may rely on this same kind of heuristic. Indeed, "additive area" explains children's area judgments better than true, mathematical area. Second, we show that children's use of "additive area" biases number judgments. Finally, to isolate "additive area" from number, we test children's area perception in a task where number is held constant across all trials. We find something surprising: even when there is no overall effect of "additive area" or "mathematical area," individual children adopt and stick to specific strategies throughout the task. In other words, some children appear to rely on "additive area," while others appear to rely on true, mathematical area – a pattern of results that may be best explained by a misunderstanding about the concept of cumulative area. We discuss how these findings raise both theoretical and practical challenges of studying quantity perception in young children.

KEYWORDS
additive area, approximate number system, area, development, number

1 | INTRODUCTION

Imagine foraging for berries. Say you come across two different kinds of berry bushes with berries that vary in shape and volume, but you are forced to forage from only one bush. Which bush do you choose to maximize the total quantity of berries? Quantity estimation tasks like this one are a ubiquitous part of everyday life (say, when selecting vegetables in a market, or estimating the size of a crowd), and successfully solving them would confer an obvious adaptive advantage. But how exactly do we solve this sort of quantity estimation task – and how does this capacity develop?

A large body of work has emphasized humans' ability to approximate visual number (Dehaene, 1997; Feigenson et al., 2004; Xu & Spelke, 2000; Xu et al., 2005), an ability that is said to be shared with some of our closest evolutionary relatives (e.g., Cantlon & Brannon, 2007). Yet, number is only one dimension of quantity estimation; relatively understudied in comparison is area (and/or volume) perception, although in many contexts this would seem to be the more relevant dimension. For example, a forager would likely seek the more *voluminous* sets of berries rather than the more *numerous* sets. Surprisingly, though, area is often seen as little more than a pesky confound in number approximation tasks (but see Barth, 2008; Brannon et al., 2006; Leibovich & Henik, 2014; Lourenco et al., 2012; Odic et al., 2013).

Here we investigate two aspects of area perception in children. First, do children, like adults, rely on a particular heuristic to estimate area in multi-element scenes? Second, do children utilize this heuristic even when the numerosity of the displays is held constant? Thus, our results bear not only on the development of area perception in its own right, but also on the way in which cumulative area influences number perception.

Research Highlights

- An "Additive-Area Heuristic" explains children's area judgments better than reliance on true, mathematical area, consistent with adult work
- "Additive area" influences children's number judgments
- In constrained tasks with controlled stimuli, children adopt specific strategies, complicating the study of area perception in children
- These results raise questions about the relative acuity for area/number, and the extent to which these dimensions have been properly isolated in prior studies

1.1 | Area perception in adults

A body of work dating back over half a century has documented an illusion of area perception: that larger objects appear relatively less large than they should. This results in underestimates of true (i.e., mathematical) area. This distortion is traditionally explained by appeal to perceptual scaling, whereby area is "scaled down," such that perceived area is equal to true area raised to some exponent less than one (Ekman & Junge, 1961; Stevens & Guirao, 1963; Teghtsoonian, 1965). Despite providing a good fit to the data, "scaling models" fall short of providing a meaningful mechanistic account of area perception. It has been suggested, for example, that these theories are ultimately unfalsifiable (insofar as they do not commit to exactly how, or to what degree, area perception is scaled; see Yousif et al., 2020). These scaling models also fail to explain some instances where two identical objects are perceived to be different in size, depending on their orientation (as scaling models should straightforwardly predict that two things equal in area are perceived as equal; see Carbon, 2016; see also Yousif et al., 2020).

In part to address these limitations in prior models, recent work has offered a (computationally) simpler explanation: that perceived area is roughly equal to the **sum** of the lengths of an object's horizontal and vertical axes. Thus, it has been suggested that area perception in adults relies on an "additive area heuristic" (so named because the visual system is ostensibly adding the dimensions of space together rather than multiplying them; Yousif & Keil, 2019, 2021a).

To understand the "additive area heuristic, consider Figure 1. Which set of dots has more cumulative area? Most participants indicate that the left panel has more area, while, in fact, the two are equal (i.e., they are equal in true, mathematical area). Notably, however, the sum of the diameters in the left panel is greater than the sum of the diameters in the right panel – and, here, the sum of the diameters is (proportionally) equal to the sum of the horizontal and vertical axes. The "additive area" account predicts that the displays with more "cumulative diameter" should be perceived as having more cumulative area. This heuristic results in fairly substantial distortions of perceived area in practice: Although adults readily discriminate between sets of dots that vary in "additive area" (while true, mathematical area is held constant), they completely fail to discriminate between sets of dots that vary in true, mathematical area (while additive area is held constant) – even in cases where mathematical area differs by as much as 30%. Adults' inability to discriminate two displays that are equated in additive area provides strong evidence that additive area is tightly linked with adults' impressions of area.

1.2 | Area perception in children

Much of our current understanding of number and area perception stems from developmental work (e.g., Brannon et al., 2006; Clearfield & Mix, 1999; Lourenco et al., 2012; Odic et al., 2013). However, area is vastly understudied relative to number. One classic study of area perception with single objects suggests that children use a "length plus width" rule for purposes of single-object area perception – a rule that resembles the additive area heuristic (Anderson & Cuneo, 1978). However, it remains unknown whether this heuristic applies to cumulative area perception. Some more contemporary studies have studied area perception directly, but often only as a comparison to number perception (Brannon et al., 2006; Lourenco et al., 2012; Odic et al., 2013). Little is known about area perception in the early years of life, especially for the sorts of stimuli canonically used to assess number perception (i.e., arrays of dots).

Understanding how children perceive area is of both theoretical and practical importance. On a theoretical level, understanding how children perceive area would speak to visual development more generally. The term "heuristic" implies a rule, possibly one that is learned through experience. Thus, children may initially perceive cumulative area veridically (i.e., by computing mathematical area) and acquire the AA heuristic that in turn leads to illusory area judgments. On a practical level, understanding area perception is necessary to guide how we measure and manipulate area in other studies of quantity perception. In general, studies of number perception make claims about number per se by controlling other dimensions. But if area perception is illusory, how should we control area perception in number displays?

One possibility is that, like adults, children's area discriminations are better explained by appeal to additive area than mathematical area. This would suggest that this tendency to use an AAH is early

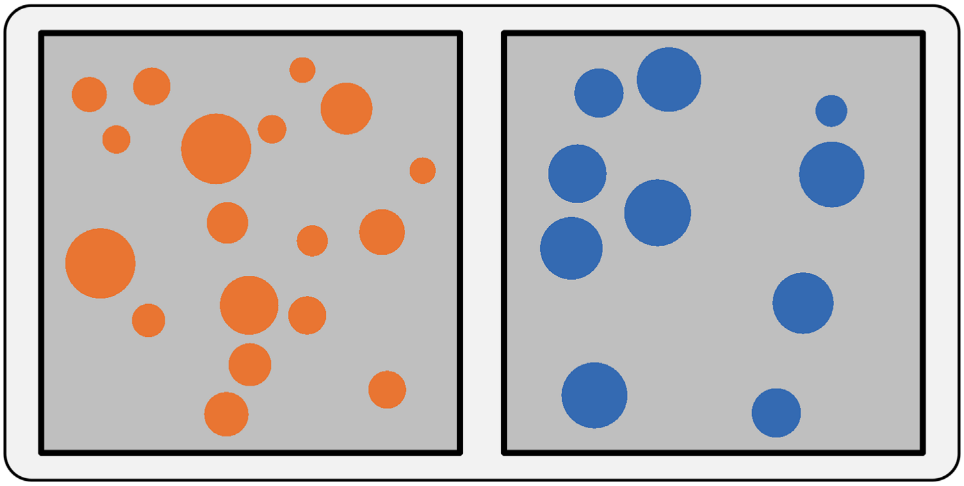

FIGURE 1 A depiction of the illusion of area, using actual stimuli from the Experiment 1. Here, most participants perceive the left panel as having more cumulative area, but in fact the two are identical. The left panel is perceived as more because it has more “additive area”

developing. Another possibility is that children are actually *better* than adults at approximating true area and *do not* use a heuristic to estimate area. If true, this would suggest that the tendency to use an AAH is one that is derived from experience, perhaps a learned rule that is successful because it provides a “shortcut to the truth.” A final possibility is that children prefer neither additive area nor mathematical area – that they haphazardly, or even strategically, rely on one cue or the other.

There are three reasons why we believe developmental data are central to this investigation. (1) While adults may rely on an AAH, children might not estimate area in the same way; knowing the answer to this question is critical to understanding one of our most foundational perceptual abilities (i.e., the ability to approximate how much “stuff” is out there). (2) Although a vast literature on approximate number has studied these kinds of estimations before, no study on quantity perception has ever asked the simplest question of all: do children perceive area veridically? (3) This approach may shed light on confounds in many existing studies, raising both theoretical and practical questions about how we ought to control visual stimuli in studies like these. For example, recent work has shown that number biases children’s area judgments (Tomlinson et al., 2020). Yet a reanalysis of those data accounting for additive area suggested this may not be the case. Indeed, when accounting for additive area, the opposite may be true: area biases number judgments (Aulet & Lourenco, 2021).

1.3 | Current study

In a first experiment, we show that children as young as 4 years do in fact rely on an AAH. In a design reminiscent of prior work in adults (e.g., Yousif & Keil, 2019), children can discriminate between stimuli which differ in additive area but not mathematical area (while the other dimension is equated). This establishes a ground truth: perceived area is not equal to true area. In a second experiment, we ask whether this insight may concretely impact studies on approximate number (given that the study of number accounts for most of the research on this topic). We show that apparent number acuity varies to a large extent depending on whether additive area or mathematical area is controlled. This finding raises questions about what “true” acuity is, and whether number is ever perceived independently of continuous spatial dimensions. In a final experiment, we test whether children still discriminate displays based on AA when specifically controlling number (borrowing stimuli from adult work which used an identical design; Yousif et al., 2020). Here, we show that individual children adopt task-specific strategies: some almost exclusively select stimuli greater in additive area, while others almost exclusively select stimuli greater in mathematical area. We discuss how children’s limited understanding of the concept of cumulative area therefore poses a concrete challenge to the study of area perception early in development.

We focus on children between the ages of 4 and 7, for three reasons. First, we purposefully wanted to cast a wide net to see if and when children begin using an AAH. Second, this age range corresponds to when children are first encountering mathematical concepts via formal schooling – and these mathematical concepts are thought to be intrinsically related to quantity estimation (Feigenson et al., 2004; Lourenco & Bonny, 2017). Finally, this age range roughly corresponds with the age range studied in the most comprehensive investigation of area approximation in children to date (Odic et al., 2013).

2 | EXPERIMENT 1: DO CHILDREN USE AN “ADDITIVE-AREA HEURISTIC”?

Do children rely on an AAH to estimate visual area? Here, children completed a forced-choice area discrimination task. The stimulus pairs were designed such that they varied either in additive area or mathematical area while the other dimension was held constant (reminiscent of the design from Yousif & Keil, 2019). The number of elements was allowed to vary freely. If children do use an AAH, we would expect that they should be able to discriminate stimuli which vary in additive area when mathematical area is equated but not stimuli which vary in mathematical area while additive area is equated.

2.1 | Method

This experiment was modeled after a previous study with adults (Yousif & Keil, 2019). As such, the stimuli in the current study closely resemble those from this prior work, though they were adapted to be more child friendly. Additional details about the stimuli can be found here: (https://osf.io/xygns/).

2.1.1 | Participants

One hundred children aged 4–7 participated (25 of 4-year-olds; 25 of 5-year-olds; 25 of 6-year-olds; and 25 of 7-year-olds). Four children who failed to complete the task were excluded prior to analysis. The study was run in a local museum. Children completed the task in a quiet area so that they were not distracted. The same children also completed Experiment 2 (with a few exceptions; see below), in a counterbalanced order (i.e., half the participants completed Experiment 1 first; the other half completed Experiment 2 first). All participants consented prior to participation, and these studies were approved by the IRB at Yale University.

2.1.2 | Materials

All of the stimuli were generated via custom software written in Python with the PsychoPy libraries (Peirce et al., 2019). The aim was to create pairs of dot array stimuli that varied in either additive area (AA) or mathematical area (MA) while the other dimension was held constant between the two arrays. For each stimulus pair, we randomly generated an initial set of discs (ranging from 20 to 100 pixels in diameter, with a buffer of at least 10 pixels between any two discs), then pseudo-randomly generated a second set of objects based on a given AA ratio (see Yousif & Keil, 2019 for more information on this process). The initial set of objects always had 10 discs. For the details of how AA, MA, and number covaried, see the "Stimulus Details" files on the OSF site linked earlier. The images depicted in Figure 1 are representative, as they were actual images used in this experiment.

In this experiment, there were only two constraints: AA and MA. There were pairs where both AA and MA were equated (to serve as a baseline), cases where MA varied while AA was controlled, and cases where AA varied while MA was controlled. While AA was controlled, area could vary between the two stimuli in either a 1.00, 1.10, 1.20, or 1.30 ratio (and vice versa for AA while MA was controlled). Because of the pseudo-random nature of stimuli creation and mathematical constraints, MA was never perfectly matched with the stated ratio; it could vary $\pm 1\%$. That is, if the MA ratio for a given trial was 1.10, then we allowed the difference in MA to fluctuate between 1.09 and 1.11. This decision was arbitrary. We could have fixed the MA ratios and allowed the AA ratios to vary slightly instead. Either way, this should not meaningfully impact the design nor the results.

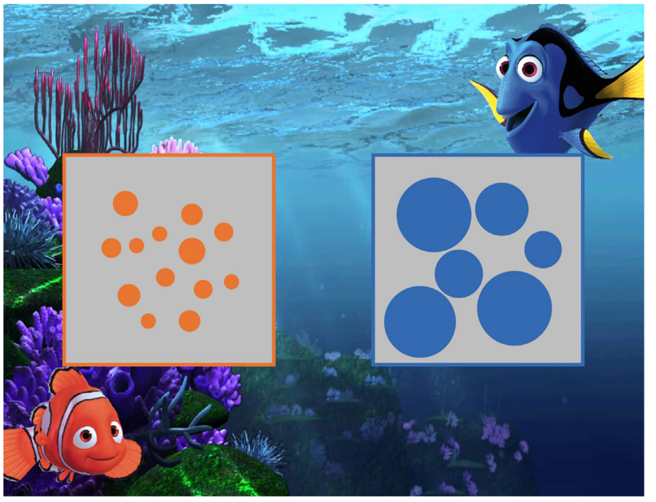

FIGURE 2 An example display, with dot displays from one of two practice stimuli that all participants saw. The practice stimuli were specifically designed to tease apart area and number. For example, here, orange is clearly greater in number, and blue is clearly greater in area. The characters were always in the same position, but "jumped" via a short animation when their side was selected. Different sets of dots appeared on each trial (unique for each experiment)

The stimuli were generated such that one set of discs was orange, and the other was blue. The orange set always appeared on the left side of the screen, and the blue set always appeared on the right side of the screen (for more information, see Section 2.1.3). The colors were fully counterbalanced with other stimulus details.

2.1.3 | Procedure

The task itself was administered on a Surface Pro 3 tablet computer using custom software written in Python. On each trial, participants saw two spatially separated sets of orange (left) and blue (right) discs. The discs themselves were on a grey background in a 400-pixel by 400-pixel frame. Those two frames were on top of an ocean scene background. Next to each stimulus was a character: Nemo was next to the orange discs, and Dory was next to the blue discs (see Figure 2). Children were told that Nemo and Dory each had painted some bubbles, and that we needed to help them figure out who had used more paint. We showed two practice examples that clearly dissociated area and number and explicitly communicated that we were asking which had more area and *not* which had more number (see Figure 2 for an example practice trial). For example, if a child chose a display with more number in the practice trials, we would say, "Well, if we were to count the bubbles, we would find that Nemo has more bubbles. But Dory's bubbles take up more space, so we should pick Dory." We would otherwise clarify as needed before continuing onto the real trials. Data would be discarded if a child demonstrated a clear failure to understand the distinction during the practice (see exclusions, above).

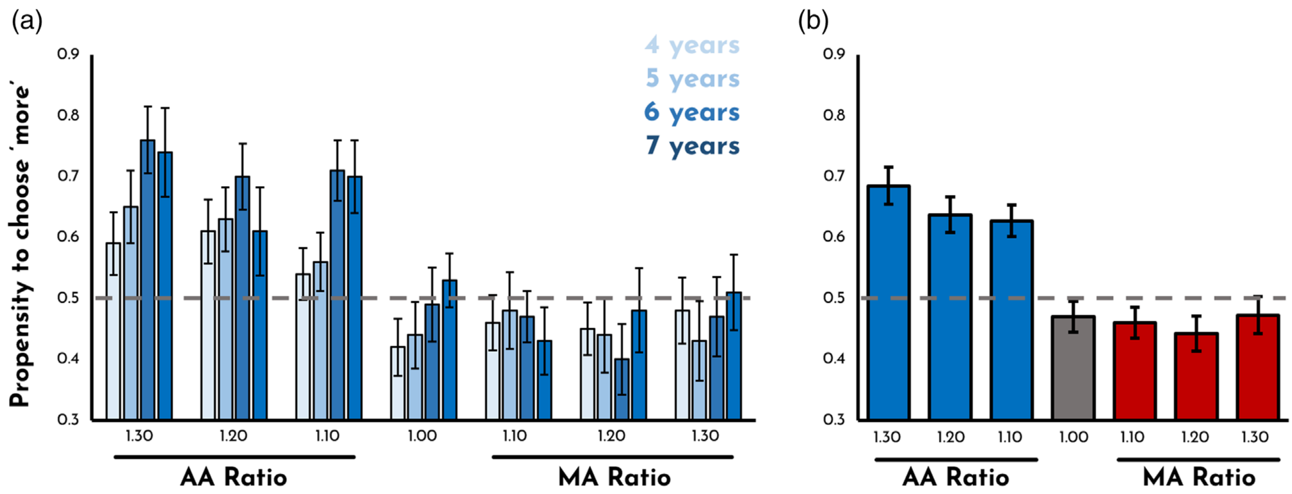

FIGURE 3 Results of Experiment 1. (a) The propensity to choose "more" (whether AA or MA) on the y-axis, broken down by ratio (see x-axis) and age (see legend in top-right). (b) The same information, but collapsed across age groups. Dashed gray lines represent chance performance. Error bars represent ± 1 SE. "Additive Area Ratio" (or AA Ratio) indicates trials which varied in AA while MA was held constant; "Mathematical Area Ratio" (or MA Ratio) indicates trials which varied in MA while AA was held constant

Children were asked to indicate their answers by saying either Nemo/Dory *or* orange/blue – whichever they felt more comfortable with. We also accepted pointing as an answer, so long as the pointing was decisive. If a child took more than a few moments to respond, we would repeat the question: "Who/which has more: [Nemo/Dory or orange/blue]?" If a child still did not respond, we would say "Nemo and Dory really need your help; can you help Nemo and Dory figure out who has more?" Occasionally, we would also ask "Who has/used more paint?" Positive reinforcement was provided anytime an answer was given, but otherwise no feedback was provided. There was also a brief animation of the selected character bouncing up and down.

Because over half of the trials had no objectively correct answer (because MA did not vary), we measured accuracy as a *propensity to choose "more"* – whether that be more AA or more MA. The side with the "correct" image was counterbalanced such that half the time the left side had more area and half the time the right side had more area. The stimuli stayed on the screen until the child indicated a choice and the experimenter submitted the response. Between each trial, there was a 1000 ms ITI. Participants completed 28 trials (7 ratios [trials where AA varied by 1.10, 1.20, and 1.30 while MA was equated; MA varying by 1.10, 1.20, and 1.30 while AA was equated; and trials where both were equated] × 2 sides [left, right] × 2 stimuli [different stimuli with identical parameters]). Trial order was randomized.

2.2 | Results and discussion

The results are shown in Figure 3. As is evident from the figure, average accuracy was above-chance for discriminations between sets which varied in AA but not MA. A repeated-measures ANOVA with two factors (condition: AA vs. MA; ratio: 1.10, 1.20, and 1.30) conducted on all participants revealed a main effect of condition ($F[1,99] = 26.96$, $p < 0.001, \eta^2 = 0.21$) but not of ratio ($F[2,98] = 1.84, p = 0.16, \eta^2 = 0.02$), and no interaction between the two ($F[2,98] = .48, p = 0.62, \eta^2 = 0.01$). Post hoc tests revealed that, overall, children were above-chance for AA discriminations ($M = 0.65$, $SD = 0.22$; $t(99) = 6.69$, $p < 0.001$, $d = 0.67$) but were slightly *below* chance for MA discriminations ($M = 0.46, SD = 0.20; t(99) = 2.03, p = 0.05, d = 0.20$).[1] Collapsing across ratios, children were consistently above-chance for AA discriminations (4-year-olds: $M = 0.58$, $SD = 0.18$; $t(99) = 2.06$, $p = 0.05$, $d = 0.41$; 5-year-olds: $M = 0.62$, $SD = 0.21$; $t(99) = 2.86$, $p = 0.009$, $d = 0.57$; 6-year-olds: $M = 0.72$, $SD = 0.18$; $t(99) = 6.16$, $p < 0.001$, $d = 1.23$; 7-year-olds: $M = 0.68$, $SD = 0.29$; $t(99) = 3.14$, $p = 0.004$, $d = 0.63$) and were never above chance for MA discriminations (in fact, performance was below chance for every group, but not significantly; 4-year-olds: $M = 0.46$, $SD = 0.15$; $t(99) = 1.24$, $p = 0.23$, $d = 0.25$; 5-year-olds: $M = 0.45$, $SD = 0.21$; $t(99) = 1.21$, $p = 0.24$, $d = 0.24$; 6-year-olds: $M = 0.45$, $SD = 0.21$; $t(99) = 1.25$, $p = 0.22$, $d = 0.25$; 7-year-olds: $M = 0.47$, $SD = 0.25$; $t(99) = 0.53$, $p = 0.60$, $d = 0.11$).

Following the example in prior work (Yousif & Keil, 2019), we also analyzed the unique contribution of number, AA, and MA. However, unlike in the adult work, there are not enough stimuli to sufficiently tease apart number and AA; the two are highly correlated. Nevertheless, a regression model would be able to reveal if these effects are driven by changes in number ratio rather than changes in AA ratio. That was not the case. A model that includes all three parameters (number, AA, and MA) is a significant predictor of participant responses ($F[3,24] = 15.55$; $p < 0.001$; $R^2 = 0.62$), yet no single cue uniquely predicts responses ($ps > 0.25$). In other words, this experiment is unable to fully dissociate whether these effects are caused by number or AA. Note, however, that this pattern of results would be no less surprising if they were driven by number and not AA: it is still mysterious that children (like adults) failed to discriminate differences in MA *as large as 30%*. A confound with number does not explain children's failure to perceive true area. To address this confound more directly, Experiment

3 uses a stimulus set with a set quantity of six items throughout the entire task.

Older children were better able to discriminate AA compared to younger children ($t(99) = 2.18$; $p = 0.031$), but were no different in their performance on MA trials ($t(99) = 0.15$; $p = 0.885$). This increased performance with age may reflect increased use of or reliance on AA, or it may reflect domain-general improvements in attention or task-taking ability. Because of this, we are reluctant to make any general claims about the developmental trends, beyond highlighting the fact that the use of MA neither increased nor decreased with age (further validating the notion that it plays little to no role in area perception).

In short, children as young as four appear to discriminate more readily on the basis of AA rather than MA. Though area acuity improves throughout development (see also Odic et al., 2013), there is no developmental change in the use of MA. Importantly, the tendency to rely on AA appears to be stable. This pattern suggests that the AAH is a foundational way of approximating area, evident even at the earliest ages we were able to test.

3 | EXPERIMENT 2: EFFECTS ON NUMBER ACUITY

What does children's use of an AAH mean for existing work? Here, we examine a single "case study": the relative acuity of number and area perception. Although prior work makes claims about the different developmental trajectories of number and area representations (Lourenco & Bonny, 2017; Odic et al., 2013), that work does not account for *perceived area*. Here, we ask whether controlling AA influences apparent number acuity. Our study is not meant to assess whether additive area *always* or *irresistibly* influences number perception, but rather to assess if it ever does. Therefore, we aim to investigate this question in the simplest way possible: by having children complete a number discrimination task with stimuli that are controlled either for additive area or mathematical area.

3.1 | Method

This experiment was identical to Experiment 1, except as indicated below. A total of 98 children aged 4–7 participated (23 of 4-year-olds, and 25 of 5-, 6-, and 7-year-olds). Note that the participants in this experiment were mostly the same as those in Experiment 1, excluding a few who completed only one portion of the task (2 of 4-year-olds who only completed the area task, and 2 of 6-year-olds who each completed only one portion). The participants who completed both experiments did so in a counter-balanced order.

Children in this experiment indicated which of the stimuli appeared greater in number. As in Experiment 1, there were two practice trials that explicitly dissociated area and number. During these trials, children were instructed to select the image with more circles, but *without counting*. If a child made the incorrect choice, we would correct them by saying "Why don't we count them? While Dory's bubbles do take up more space, we can see that Nemo actually has more bubbles, so we should pick Nemo. We want to do the same thing next time, but without counting." If a child continued to count in later trials, we would discourage them from doing so. For example, we might say that Nemo and Dory needed our help quickly and that we did not have time to count.

This experiment used an independent set of stimuli. The default stimulus always had 10 discs. The second stimulus had either 11, 13, 15, or 17. Half the trials were AA-controlled (while MA varied randomly; the average MA ratio was 0.79, with a minimum of 0.62 and maximum of 0.95), and the other half were MA-controlled (while AA varied randomly; the average AA ratio was 1.15, with a minimum of 1.02 and a maximum of 1.31). There were 32 trials total (4 number ratios [1.10, 1.30, 1.50, 1.70] × 2 trial types [AA-controlled, MA-controlled], × 2 sides [left, right] × 2 stimuli [different stimuli with identical parameters]). Full stimulus details are available on the OSF page.

3.2 | Results and discussion

First, we checked to ensure that participants were successfully completing the number discrimination task. Across all age groups, participants were above-chance at making number discriminations, regardless of whether the stimuli were controlled for AA ($M = 0.71$, $SD = 0.16$; $t(97) = 13.50$, $p < 0.001$, $d = 1.36$) or MA ($M = 0.77$, $SD = 0.17$; $t(99) = 16.04$, $p < 0.001$, $d = 1.62$). This was independently true for all ages ($ps < 0.001$; $ds > 1.02$). Furthermore, collapsed across trial type (AA- vs. MA-controlled) there was a clear effect of ratio such that participants were better able to discriminate displays with a greater difference in number ($F(3, 291) = 46.32$, $p < 0.001$, $\eta^2 = 0.32$). For AA-controlled trials, accuracy ranged from 0.52 for the lowest ratio to 0.78 for the highest ratio (but note that the second highest ratio, 1.5, had a slightly higher accuracy of 0.80); for MA-controlled trials, accuracy ranged from 0.68 for the lowest ratio to 0.84 for the highest ratio. However, the critical question was whether participants' number acuity differed depending on how area was controlled; these results are shown in Figure 4. For all age groups (collapsed across ratio), number acuity was *higher* when MA was controlled than when AA was controlled. This pattern was insignificant for the 4- and 5-year-olds ($ps > 0.14$; $ds > 0.18$), but significant for the 6- ($t(24) = 2.34$, $p = 0.03$, $d = 0.47$) and 7-year-olds ($t(24) = 3.19$, $p < 0.005$, $d = 0.64$). This finding is consistent with the results of adult studies indicating similar variation in acuity (Yousif & Keil, 2020). As seen in Figure 4, though, these effects appear to be driven by the smallest and largest ratios (1.1 and 1.7). This is true: collapsed across age groups, there was a significant difference for the 1.1 ($p < 0.001$) and 1.7 ($p = 0.032$) ratios, but no significant difference for the 1.3 or 1.5 ratios ($ps > 0.750$). It is unclear why this effect is driven by the smallest and largest ratios tested. This may be due to task-specific strategies that differ across ratios (see Experiment 3), but we wish to refrain from over-interpreting this pattern of results for now.

These results raise questions about how individual differences in acuity should be interpreted (e.g., as in Cordes & Brannon, 2008, 2009; Odic et al., 2013). In other words, under which condition (i.e., AA control or MA control) is performance a better indicator of "true" number

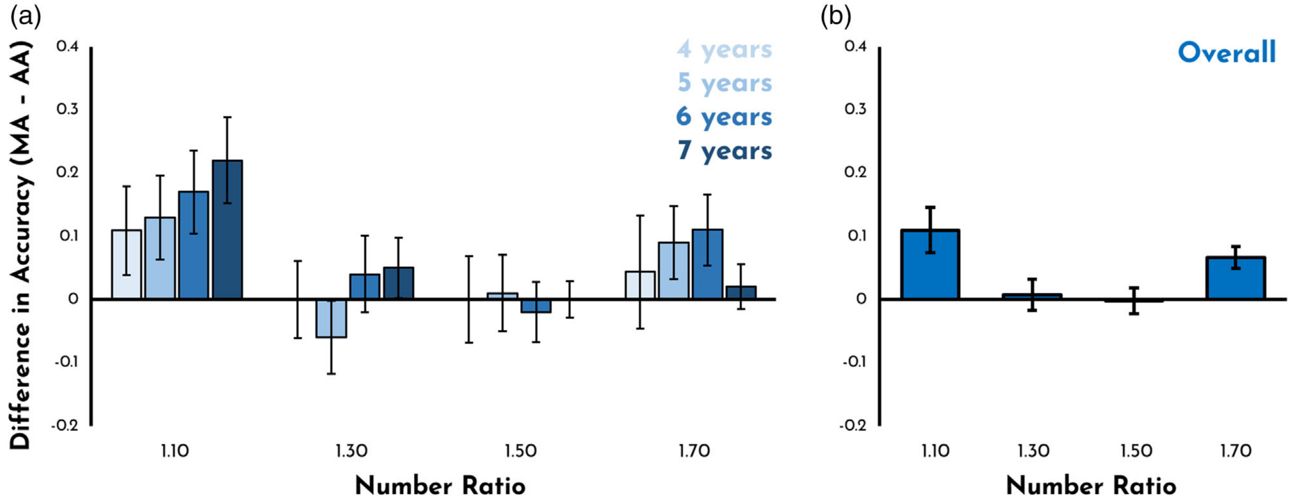

FIGURE 4 Results of Experiment 2. (a) Difference in accuracy (MA-AA) is plotted on the y-axis, broken down by ratio (x-axis) and age (see legend in top-right). (b) The same information but collapsed across age groups. Positive values indicate better performance when MA is controlled; negative values indicate better performance when AA is controlled. Error bars represent ± 1 SE

discrimination ability? Moreover, how should relative differences in acuity be interpreted, especially in light of correlations with each other and other cognitive abilities (e.g., Lourenco & Bonny, 2017; Lourenco et al., 2012; Odic et al., 2013)? We do not intend to imply that "true" acuity can never be discerned; instead, we only wish to highlight a potential complexity in comparing relative acuity so that future work may address it head-on.

Relatedly, a large body of work has concerned itself with "congruity effects" between number and other continuous magnitudes (e.g., Brannon et al., 2004; Hurewitz et al., 2006; Rousselle et al., 2004). For example, it has been suggested that representations of time, space, quantity, and other magnitudes rely on similar cortical processes, or one "general magnitude" representation (Lourenco & Longo, 2010; Sokolowski et al., 2017; Walsh, 2003). In support of this view, prior work has identified Stroop-like errors between area and number (Brannon et al., 2004; Hurewitz et al., 2006; Rousselle et al., 2004) – much like those observed here. Our results bear on these past findings in two key ways: (1) they demonstrate a congruity effect for a novel dimension (i.e., AA, which has never been studied in the context of congruity effects) and (2) they raise questions about the *cause* of previously observed congruity effects. As Yousif and Keil (2019) argue, it is possible that this confound between AA (which has not been studied in this context) and MA (which is often studied in this context) may lead to the appearance of a bi-directional congruity effect between area and number. Consistent with this possibly, in Experiment 1, we provide evidence that number seems to have little or no influence on area estimates. Yet, here, we have provided evidence that perceived area seems to influence number judgments.

4 | EXPERIMENT 3: EQUATING NUMBER

How can we be sure that the results of Experiment 1 implicate "additive area" specifically? While much work has addressed the relation between number and area indirectly (see, e.g., Yousif & Keil, 2019, 2020), recent work with adults has demonstrated an effect of additive area on area judgments even in stimuli that are perfectly equated for number (Yousif et al., 2020). In other words, even when number was set at a fixed value (e.g., 6 or 10) across *all* trials, adult participants still discriminated displays only based on "additive area." Here, we borrow this exact design: we use a subset of the stimuli tested on adults and ask whether children, like adults, discriminate visual displays on the basis of additive area even when number is controlled.

4.1 | Method

This experiment was identical to Experiment 1 except as noted. Fifty children (25 of 4-year-olds and 25 of 5-year-olds) participated. We specifically targeted the youngest age ranges tested in our previous experiments (4- and 5-year-olds). Because these youngest children demonstrated use of additive area in Experiment 1, showing that they use additive area in this number-controlled task would provide a relatively definitive demonstration that children do in fact robustly rely on this cue (independently from other factors). Due to the COVID-19 pandemic, approximately half of our data were collected in-person (14 of 4-year-olds and 10 of 5-year-olds) and half were collected online (11 of 4-year-olds and 15 of 5-year-olds). Eleven additional children failed to complete the task, and so were excluded prior to data analysis.

All the stimuli used in this experiment were borrowed from a previous study on adults' area estimation (Yousif et al., 2020). The goal of this study was to use stimuli that were equated in number. In other words, every single display that children saw throughout the experiment had exactly six dots; number was never manipulated. Children who completed the task in person completed 40 trials of four unique trial types. For half of the trials, AA varied while MA was equated (in two different ratios; 1.15 and 1.10); for the other half of trials, MA varied while

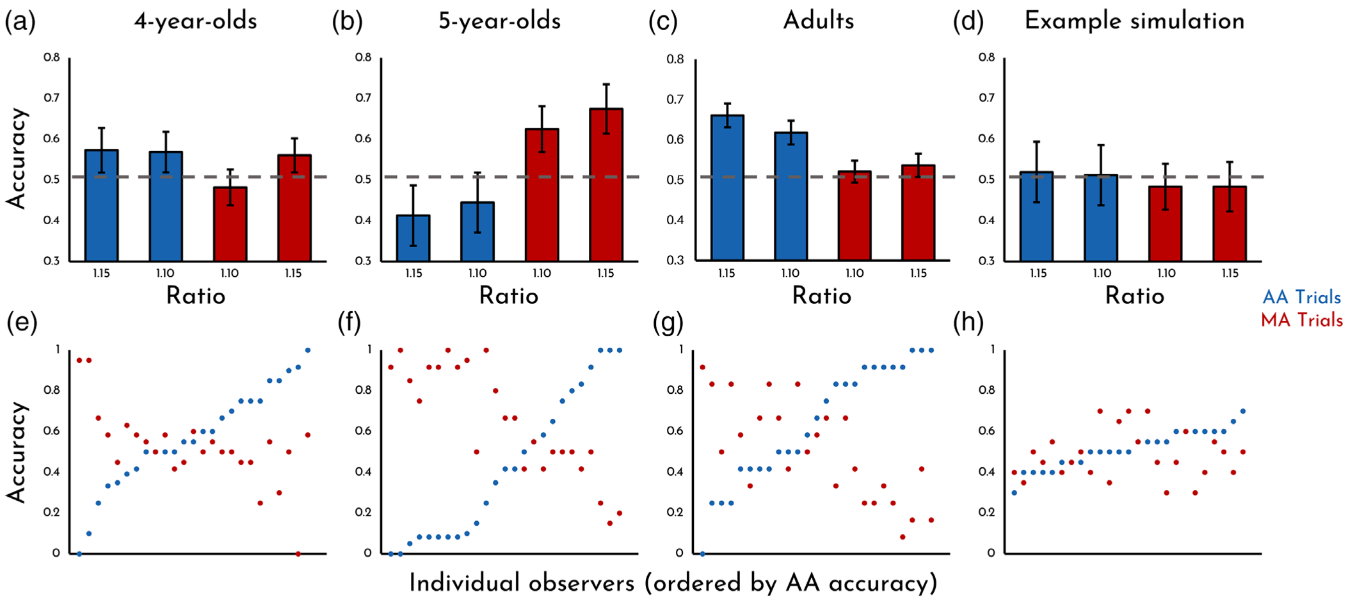

FIGURE 5 Results of Experiment 3. Overall results for four-year-olds (a), 5-year-olds, (b) and adults (c), per Yousif et al., 2020. Accuracy is depicted on the y-axis, broken down by ratio (see x-axis). (d) Simulated results for this same experiment. Dashed gray lines represent chance performance. Error bars represent ± 1 SE. Blue bars indicate trials which varied in AA while MA was held constant; red bars indicate trials which varied in MA while AA was held constant. (e–h) Accuracy shown for each observer, for 4-year-olds (e), 5-year olds (f), and adults (g) as well as simulated data (h). The adult data in (c) and (d) has been randomly down-sampled so that the number of participants is matched to the kid data. Of note here is the striking degree of polarization: children seem to clearly rely on one cue or the other, such that performance for the two cues is strongly negatively correlated (more so than simulations of random behavior, as well as adults)

AA was equated (in the same two ratios; 1.15 and 1.10). There were 10 unique trials for each of the four trial types. To make the task more manageable for online participants, children who completed the task online completed 24 trials. The counterbalancing was identical to that of the in-person participants, except that these participants saw six unique trials for each trial type instead of 10.

4.2 | Results and discussion

The results are shown in Figure 5A and B. We first analyzed data from the 4-year-olds. Participants were unable to discriminate both AA and MA. We conducted one-sample *t*-tests on each AA/MA ratio. Participants were unable to discriminate both AA (1.10 ratio: $M = 0.57$, $SD = 0.25$, $t[24] = 1.40$, $p = 0.18$, $d = 0.28$; 1.15 ratio: $M = 0.57$, $SD = 0.27$, $t[24] = 1.34$, $p = 0.19$; $d = 0.27$) and MA (1.10 ratio: $M = 0.48$, $SD = 0.22$, $t[24] = 0.40$, $p = 0.69$, $d = 0.08$; 1.15 ratio: $M = 0.56$, $SD = 0.21$, $t[24] = 1.48$, $p = 0.15$; $d = 0.30$).

We then analyzed data from the 5-year-olds. We again conducted one-sample *t*-tests on each AA/MA ratio. Participants were unable to discriminate AA (1.10 ratio: $M = 0.45$, $SD = 0.37$, $t[24] = 0.74$, $p = 0.47$, $d = 0.15$; 1.15 ratio: $M = 0.41$, $SD = 0.37$, $t[24] = 1.17$, $p = 0.24$; $d = 0.23$) but *did* discriminate using MA (1.10 ratio: $M = 0.63$, $SD = 0.28$, $t[24] = 2.22$, $p = 0.04$, $d = 0.45$; 1.15 ratio: $M = 0.68$, $SD = 0.30$, $t[24] = 2.87$, $p = 0.008$; $d = 0.58$). In other words, unlike any pattern previously observed in adults, children successfully discriminated between stimuli that varied in true area while AA was held constant. However, these overall statistics obfuscate an important aspect of these data: that children's responses are highly polarized (as can be seen in Figure 5E and F). In other words, individual children tend to rely exclusively on AA or MA, to such an extent that suggests children may have relied on overt strategies to complete the task (rather than responding based on their genuine impression on a trial-by-trial basis). In fact (as can be seen in Figure 5G and H), children are *more* polarized than both simulated random data and adults (who completed this task on identical stimuli).

To confirm as much, we conducted permutation tests to assess what level of "polarization" (i.e., children using *only* AA or *only* MA) would be expected by random chance. We simulated one million repetitions of our experiment, each of which included simulated data from 25 participants completing 40 trials each. For each trial, we randomly generated a response. For each simulation (and for our own data) we calculated three "polarization scores": one for the overall data, one for the AA trials, and one for the MA trials. For the overall data, polarization scores were calculated as the difference between the number of AA selections and the number of MA selections. For the individual trial types, polarization scores were calculated as the number of selections over or under chance performance. So if there were eight trials, at-chance performance would involve four selections in favor of AA and four against; if a participant selected AA six times, we would assign a polarization score of 2. The code used to conduct these permutation tests, as well as the resulting data, are available on our OSF page.

The average overall polarization score in one million simulations was 2.51 ($SD = 0.39$). The average polarization scores for AA and MA in those million simulations were both 1.76 ($SD = 0.28$). First, we compared the 4-year-olds' data to these simulated values. Their overall polarization score was 4.99, a value more than six standard deviations greater than the mean value in our simulations ($p < 0.001$). Their overall polarization score for AA trials was 3.45 (again, six standard deviations greater than the mean; $p < 0.001$) and for MA trials was 3.10 (five standard deviations greater than the mean; $p < 0.001$). Next, we compared the 5-year-olds' data to these simulated values. Their overall polarization score was 8.36, a value more than 11 standard deviations greater than the mean value in our simulations ($p < 0.001$). Their overall polarization score for AA trials was 4.88 (11 standard deviations greater than the mean; $p < 0.001$) and for MA trials was 3.16 (five standard deviations greater than the mean; $p < 0.001$). Collectively, these results demonstrate that children adopted distinct strategies in this task (see Figure 5E–H for a visual depiction of the polarized responses, in comparison with simulated random data).

Unlike in Experiment 1, 4-year-olds were in this experiment unable to discriminate stimuli on the basis of AA. By age 5, children exhibit above-chance performance for MA but not AA trials. We propose that this pattern reflects children's adoption of a strategy particular to the stimuli for this task. When number is held constant between dot arrays, the only way to dissociate additive area and mathematical area is to manipulate the *variance* in dot sizes across sets. In practice, this means that the set with more mathematical area, or less additive area, will almost always contain the single largest item. Thus, lacking a complete understanding of cumulative area, one strategy that children may use is to select the single largest object, rather than the largest set, to guide their selections. Another strategy that children might use is to select the display with more uniform dot sizes. A participant adopting this strategy would almost always choose the display with more AA, or with less MA (and so, again, could explain the polarization that we observe). Both of these strategies are reflected in the polarized responses that we observed.

Another way to think about the difference in performance between children and adults is a difference in local versus global perception (or local vs. global attention). Whereas adults may more readily attend to the "ensemble" of dots (e.g., Marchant et al., 2013), children may more readily attend to local features (i.e., the largest individual dot). That said, even children perceive ensembles in some cases, and their average size representations, like adults, are based on *diameter* not cumulative area (Sweeny et al., 2015) – exactly as the "additive area heuristic" would predict. It would be strange indeed if children's *average* size estimations were based on an "additive heuristic," but that their *cumulative* size estimations were based on true, mathematical area. This is more reason to think that additive area provides a more general explanation of size perception across tasks.

We would draw analogy to classic Piagetian tasks, in which children exhibit some failures that are not likely due to *perception* but instead due to some failure to understand the task. For example, why do children say that a set of pennies that is spread out is more numerous than an (equal) set of pennies that is less spread out (Piaget, 1965)? Surely children do not *see* more pennies in one case than the other. Indeed, it has been shown that when the question is framed differently, children understand that the two sets are equal in number (Hudson, 1983). Similarly, here, we think the most conservative interpretation of these data is that, when faced with highly constrained stimuli, children will fall back on task-specific strategies. We cannot definitively conclude that children use either AA or MA for purposes of area estimation.

5 | GENERAL DISCUSSION

In Experiment 1, we showed that children, like adults, relied on an additive area heuristic to estimate area. In Experiment 2, we showed how this insight may challenge our current understanding of the relationship between number and area perception, raising questions about what ought to be considered "true" acuity, and whether prior studies have properly isolated number. Finally, in Experiment 3, we showed that when stimuli are constrained and ratios are small, children fall back on specific strategies to make area judgments. These findings preliminarily suggest that children, like adults, may rely on a heuristic to estimate visual area; however, these findings also highlight a number of theoretical and practical challenges that prevent any strong claims about children's area perception (and, consequently, number perception).

Isolating number from other continuous spatial dimensions is necessary to demonstrate that children and infants can truly estimate approximate number. Thus, it is not surprising that this concern has been raised before. A popular counterpoint to some of the initial work positing large number discrimination in infancy (Xu & Spelke, 2000) was that infants rely on contour length (or perimeter) rather than number to discriminate between dot displays that vary in number (Clearfield & Mix, 1999). However, this point was dismissed when later work accounted for this confound (Xu et al., 2005), and when new manipulations began controlling contour length in clever ways (e.g., McCrink & Wynn, 2007). Since then, it has been generally accepted that children and infants can discriminate number – all while area estimation itself has been largely ignored.

It is important to consider exactly *how* number, mathematical area, and additive area are confounded in dot displays like those used here. Critically, controlling for mathematical area while manipulating number (as many studies do; e.g., Xu & Spelke, 2000) actively *creates* a confound with additive area. Imagine two sets of dots which are equated for MA but vary in number, such that one has 30% more number than the other. All else equal, the display with more number will *also* have more additive area (about 15% more, on average), and thus more perceived area. In other words: *perceived* area and number are still confounded - posing a serious challenge to the interpretation of many studies, and numerous practical challenges for stimulus creation in both adults and children (for a lengthier explanation of this issue, see also Yousif & Keil, 2021a).

In the contemporary quantity perception literature, a few other studies have investigated area estimation explicitly (Brannon et al., 2006; Lourenco et al., 2012; Odic et al., 2013). For example, it has been suggested that children of all ages have greater area acuity than number acuity (Odic et al., 2013; but for contradictory results in infants, see

Cordes & Brannon, 2008, 2009). The present work does not directly challenge this finding, but it does challenge its interpretation. For example, number acuity varies across displays that control for either additive area or mathematical area – but which of these area estimates reflects true acuity? It is possible, in theory, that number acuity is equal to or greater than area acuity throughout development, but that existing studies failed to properly measure true number acuity, area acuity, or both. Further, work with adults suggests that additive area dramatically influences the perception of number (but that number does not appear to influence the perception of area; see Yousif & Keil, 2020), raising the concern that number estimation cannot be isolated at all (and that there is no "true" area or number acuity in the first place; see also Leibovich et al., 2017). Our view is that studies on quantity perception should more seriously consider the role that perceptual heuristics may play in children's judgments.

5.1 | Reasons to believe children *do* rely on an "additive area heuristic"

The most crucial finding here is not about number perception, but area perception: the present results (in combination with work in adults; Yousif & Keil, 2019, 2020; Yousif et al., 2020) provide compelling evidence that impressions of area do not reflect reality (even if the precise mechanisms remain uncertain). Here, we have proposed a specific rule that governs area perception: that children (like adults) *add* the dimensions of space together rather than multiply them. Experiment 1 provided evidence in support of this hypothesis: children discriminated displays that differed in additive area but failed to discriminate displays that were equated in terms of additive area.

However, regression analyses revealed that neither additive area nor number was a significant predictor of responses in Experiment 1 – raising the possibility that the patterns observed here are not about area perception at all, but, instead, reflect a conflation with number. Yet there are several reasons to believe that additive area nevertheless provides the best explanation. First, note that this study was not designed to tease apart additive area from number; this would require significantly more data. Second, the regression analyses are imperfect here, given that more than half of the stimuli do not vary in additive area (and so detecting variance based on additive area is more challenging). Third, although number may explain a tendency to choose the stimuli with more additive area, it cannot explain the failure to discriminate displays that vary in true, mathematical area. If number explained these results, one may straightforwardly predict *below-chance* performance for the mathematical area trials (but this is not what we find). Fourth, area/number congruity effects are often modest – statistically discernable in large studies, but not visually apparent. There is no compelling reason to believe that congruity effects could result in such large distortions of perceived area. Finally, there is the matter of parsimony: prior work has shown numerous compelling demonstrations that something like an "additive area heuristic" explains adults' impressions of area. When we observe the same pattern in children, it may be parsimonious to assume that there is one underlying explanation (keeping in mind prior work demonstrating the use of a similar heuristic; Anderson & Cuneo, 1978).

Beyond the data here, there are many reasons that "additive area" may provide a general explanation for area perception. First, it has been shown that even exceedingly simple manipulations challenge other models of area perception. For example, simply rotating squares 45 degrees increases their perceived area (Yousif et al., 2020), a finding that is at odds with both veridical models of area perception and classic "scaling models" of area perception (e.g., Stevens & Guirao, 1963). Second, additive area better explains area judgments than many possible related dimensions (e.g., cumulative perimeter; Yousif & Keil, 2019; Yousif et al., 2020). Third, these effects have been shown to generalize across a range of 2D shapes, including circles (Yousif & Keil, 2019, 2020), rectangles (Yousif & Keil, 2019), ellipses (Yousif & Keil, 2019), diamonds (Yousif et al., 2020), and squares (Yousif et al., 2020). Fourth, and perhaps more importantly, this illusion of size perception extends beyond 2D area perception; judgments of 3D volume (for both cubes and spheres) also appear to be explained by an "additive heuristic" (Bennette et al., 2021). Finally, biases of children's area perception like these have been well-documented for decades, at least for single-object comparisons (e.g., Anderson & Cuneo, 1978). Collectively, these findings suggest that the visual system may generally fail to accurately integrate information across multiple spatial dimensions (see Carbon, 2016) – a problem that is resolved by using an "additive heuristic."

5.2 | Theoretical and practical challenges to studying area perception in children

Although our findings provide some compelling reasons to believe that children, like adults, may rely on an "additive area heuristic," they also raise numerous theoretical and practical challenges to the study of area perception in the early years of life.

(1) ***Signal Clarity Theory***. This view (Cantrell & Smith, 2013) posits that dimensions with higher variance will be more salient to participants in general than dimensions with lower variance. This sort of explanation could apply to the results of Experiment 1. Although additive area and mathematical area specifically varied within the same range (from 1.00 to 1.30), number did not. Instead, number varied in a range from 0.70 to 1.80 – so there is good reason to believe that this information, according to Signal Clarity Theory, should be more salient. Despite this, number appears not to be the predominant explanation of behavior: regression analyses clearly revealed that number alone could not explain children's responses. In other words, Signal Clarity Theory suggests that our design in Experiment 1 is biased *in favor* of number – yet number still fails to explain the patterns we observe.

(2) ***Existing models with many components***. Several proposals have been made about how to measure or control many continuous dimensions at the same time (e.g., Barth, 2008; DeWind et al., 2015; Gebuis & Reynvoet, 2011; Salti et al., 2017). However, these holistic approaches typically do not focus on *area* perception (but see Barth, 2008). Part of our suggestion is that these complex models should be applied to

area perception with the same level of rigor as models being applied to number perception. Indeed, when this has happened, hints of the present effects emerge. For example, Barth (2008) found that cumulative diameter (functionally equivalent to additive area in many stimuli) was an excellent predictor of area judgments – as good, or better, than area itself (i.e., mathematical area). DeWind et al. (2015) similarly found that the second-best predictor of number judgments was total perimeter (also functionally equivalent to additive area in many stimuli).

(3) ***Flawed stimulus design?*** A recent critique construes the effects found here, and in related work, as products of a "flawed stimulus design" (Park, 2021; but see Yousif & Keil, 2021b). The argument goes as follows: In order to dissociate additive area from mathematical area, other dimensions, like number, have to vary in unusual ways. As theories like Signal Clarity Theory point out, this could lead to a difference in salience across dimensions. This is true. As we highlighted above, number varies to a much greater degree in Experiment 1 than either additive area or mathematical area. The critical question here is whether this matters in practice. In our results, it does not: If anything, in Experiment 1 there is a bias in favor of number – and yet number is not the sole factor explaining children's judgments (for a more in-depth treatment of this argument, see Yousif & Keil, 2021b). In Experiment 3, number is not a factor at all – yet we still observe a systematic pattern of results.

There still might be interplay between dimensions and it is always important to consider how other dimensions are measured or manipulated. But we do now need to acknowledge that area perception seems to be illusory (see Yousif & Keil, 2021a, 2021b, for review). Given that this is true, we need to think deeply about how congruity effects between number and true, mathematical area ought to be interpreted – and how children come to tease these dimensions apart.

(4) ***Is this really a confound?*** Experiment 2 demonstrates a single instance in which controlling for additive area vs. mathematical area could result in a different pattern of results (but for more examples, see Yousif & Keil, 2020). While these effects are empirically modest, they are theoretically important, as entire perspectives rest on them (e.g., on the prediction that there are bi-directional congruity effects between area and number). Whether, or to what extent, additive area is confounded with number is highly dependent on the specific manner in which the stimuli are constructed, and on how area and number are controlled relative to one another. Many studies employ a method of using both area-congruent trials (in which number and area are positively correlated) and area-incongruent trials (in which number and area are negatively correlated). Yet even with very modest assumptions about how stimuli are created, AA is almost always going to be confounded with number and mathematical area (see Yousif & Keil, 2021a). (This does not mean that additive area is *necessarily* confounded, but rather that it just *happens to be* given the way these stimuli are typically designed.)

(5) **Children's understanding of cumulative area**. It is challenging to convey the idea of *cumulative area* to young children. One may wonder, for example, if children chose to sometimes select the display with the single largest object (rather than the set with the most cumulative area), as may have been the case in Experiment 3. Without any way to ensure that children understand cumulative area, we can only speculate about why they adopted the strategies that they did. However, data from Experiment 3 clearly demonstrate that children are being strategic in some way (whether because they misunderstand the question about cumulative area, or for some other reason). Thus, we think that future work should be especially careful when probing children's impressions of cumulative area.

5.3 | Conclusion

Do children estimate area using an "additive area heuristic"? These results offer support for such an account. However, we identify several practical concerns – namely, children's limited understanding of the concept of cumulative area – that prevent any strong conclusions. Nevertheless, these findings raise question about how prior studies have measured both number and area perception. Thus, the present results suggest new theoretical and practical directions for the field to consider. This serves as a first step toward understanding children's visual impressions of – as well as their conception of – cumulative area.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available here: https://osf.io/xygns/?view_only=bd54c517a15d46249c1ca72aa369251f

ENDNOTE

[1] Note that this experiment and the subsequent experiment were conducted almost entirely on the same group of children in a counterbalanced order. Therefore, one may wonder whether the order of these two experiments impacted children's performance on these tasks. In short, here, and for all analyses reported in this paper, there were no differences between the children that completed this experiment first versus the other experiment first. While we do not exhaustively report the differences between these groups for every analysis in this paper, we wanted to make note of the similarity across conditions for this analysis, which is clearly most integral to our conclusions. Children who completed this experiment *first* had an average accuracy of 0.68, 0.61, 0.61, 0.43, 0.44, 0.45, and 0.49 for the AA-1.30, AA-1.20, AA-1.10, Equal, MA-1.10, MA-1.20, and MA-1.30 ratios, respectively. Children who completed this experiment second had very similar values of 0.68, 0.67, 0.65, 0.52, 0.48, 0.43, and 0.46, respectively. The basic AA effect (AA vs. chance) is significant in both cases, $p < 0.001$.

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. https://doi.org/10.1037/0096-3445.107.4.335

Aulet, L. S., & Lourenco, S. F. (2021). The relative salience of numerical and non-numerical dimensions shifts over development: A re-analysis of Tomlinson, DeWind, and Brannon (2020). *Cognition*, *210*, 104610. https://doi.org/10.1016/j.cognition.2021.104610

Barth, H. C. (2008). Judgments of discrete and continuous quantity: An illusory Stroop effect. *Cognition*, *109*, 251–266. https://doi.org/10.1016/j.cognition.2008.09.002

Bennette, E., Keil, F. C., & Yousif, S. R. (2021). A ubiquitous illusion of volume: Are impressions of 3D volume captured by an "Additive Heuristic"? *Perception*, *50*, 462–469. https://doi.org/10.1177/03010066211003746

Brannon, E. M., Abbott, S., & Lutz, D. J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition*, *93*, 59–68. https://doi.org/10.1016/j.cognition.2004.01.004

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. https://doi.org/10.1111/j.1467-7687.2006.00530.x

Cantlon, J. F., & Brannon, E. M. (2007). How much does number matter to a monkey (*Macaca mulatta*)? *Journal of Experimental Psychology: Animal Behavior Processes*, *33*, 32–41.

Cantrell, L., & Smith, L. B. (2013). Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition*, *128*, 331–352. https://doi.org/10.1016/j.cognition.2013.04.008

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

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408–411. https://doi.org/10.1111/1467-9280.00177

Cordes, S., & Brannon, E. M. (2008). Quantitative competencies in infancy. *Developmental Science*, *11*, 803–808. https://doi.org/10.1111/j.1467-7687.2008.00770.x

Cordes, S., & Brannon, E. M. (2009). The relative salience of discrete and continuous quantity in young infants. *Developmental Science*, *12*, 453–463. https://doi.org/10.1111/j.1467-7687.2008.00781.x

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford University Press.

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. https://doi.org/10.1016/j.cognition.2015.05.016

Ekman, G., & Junge, K. (1961). Psychophysical relations in visual perception of length, area and volume. *Scandinavian Journal of Psychology*, *2*, 1–10. https://doi.org/10.1111/j.1467-9450.1961.tb01215.x

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods*, *43*, 981–986. https://doi.org/10.3758/s13428-011-0097-5

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development*, *54*, 84–90. https://doi.org/10.2307/1129864

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences*, *103*, 19599–19604. https://doi.org/10.1073/pnas.0609485103

Leibovich, T., & Henik, A. (2014). Comparing performance in discrete and continuous comparison tasks. *The Quarterly Journal of Experimental Psychology*, *67*, 899–917. https://doi.org/10.1080/17470218.2013.837940

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, *e164*, 1–62.

Lourenco, S. F., & Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Developmental Science*, *20*, e12418. https://doi.org/10.1111/desc.12418

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*, *109*, 18737–18742. https://doi.org/10.1073/pnas.1207212109

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*, *21*, 873–881. https://doi.org/10.1177/0956797610370158

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. https://doi.org/10.1016/j.actpsy.2012.11.002

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, *18*, 740–745. https://doi.org/10.1111/j.1467-9280.2007.01969.x

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. https://doi.org/10.1037/a0029472

Park, J. (2021). Flawed stimulus design in additive-area heuristic studies. https://doi.org/10.31234/osf.io/h6uw8

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y

Piaget, J. (1965). *The child's conception of number*. Norton.

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. https://doi.org/10.1016/j.jecp.2003.10.005

Salti, M., Katzin, N., Katzin, D., Leibovich, T., & Henik, A. (2017). One tamed at a time: A new approach for controlling continuous magnitudes in numerical comparison tasks. *Behavior Research Methods*, *49*, 1120–1127. https://doi.org/10.3758/s13428-016-0772-7

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. https://doi.org/10.1016/j.neuropsychologia.2017.01.019

Stevens, S. S., & Guirao, M. (1963). Subjective scaling of length and area and the matching of length to loudness and brightness. *Journal of Experimental Psychology*, *66*, 177. https://doi.org/10.1037/h0044984

Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2015). Ensemble perception of size in 4–5-year-old children. *Developmental Science*, *18*, 556–568. https://doi.org/10.1111/desc.12239

Teghtsoonian, M. (1965). The judgment of size. *The American Journal of Psychology*, *78*, 392–402. https://doi.org/10.2307/1420573

Tomlinson, R. C., DeWind, N. K., & Brannon, E. M. (2020). Number sense biases children's area judgments. *Cognition*, *204*, 104352. https://doi.org/10.1016/j.cognition.2020.104352

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, *8*, 88–101. https://doi.org/10.1111/j.1467-7687.2005.00395.x

Yousif, S., Aslin, R., & Keil, F. (2020). Judgments of spatial extent are fundamentally illusory: 'Additive-area' provides the best explanation. *Cognition*, *205*, 104439. https://doi.org/10.1016/j.cognition.2020.104439

Yousif, S. R., & Keil, F. C. (2019). The additive area heuristic: An efficient but illusory means of visual area approximation. *Psychological Science*, *30*, 495–503. https://doi.org/10.1177/0956797619831617

Yousif, S. R., & Keil, F. C. (2020). Area, not number, dominates estimates of visual quantities. *Scientific Reports*, *10*, 1–13. https://doi.org/10.1038/s41598-020-68593-z

Yousif, S. R., & Keil, F. C. (2021a). How we see area and why it matters. *Trends in Cognitive Sciences*, *25*, 554–557. https://doi.org/10.1016/j.tics.2021.03.017

Yousif, S. R., & Keil, F. (2021b). Quantity perception: The forest and the trees. https://doi.org/10.31234/osf.io/ydcg7

How to cite this article: Yousif, S. R., Alexandrov, E., Bennette, E., Aslin, R. N., & Keil, F. C. (2022). Do children estimate area using an "Additive-Area Heuristic"? *Developmental Science*, e13235. https://doi.org/10.1111/desc.13235