

# Eye movements reveal distinct encoding patterns for number and cumulative surface area in random dot arrays

Darko Odic

Department of Psychological and Brain Sciences,  
University of British Columbia, Vancouver, BC, Canada



Justin Halberda

Johns Hopkins University, Baltimore, MD, USA

Humans can quickly and intuitively represent the number of objects in a scene using visual evidence through the Approximate Number System (ANS). But the computations that support the encoding of visual number—the transformation from the retinal input into ANS representations—remain controversial. Two types of number encoding theories have been proposed: those arguing that number is encoded through a dedicated, enumeration computation, and those arguing that visual number is inferred from nonnumber specific visual features, such as surface area, density, convex hull, etc. Here, we attempt to adjudicate between these two theories by testing participants on both a number and a cumulative area task while also tracking their eye-movements. We hypothesize that if approximate number and surface area depend on distinct encoding computations, saccadic signatures should be distinct for the two tasks, even if the visual stimuli are identical. Consistent with this hypothesis, we find that discriminating number versus cumulative area modulates both *where* participants look (i.e., participants spend more time looking at the more numerous set in the number task and the larger set in the cumulative area task), and *how* participants look (i.e., cumulative area encoding shows fewer, longer saccades, while number encoding shows many short saccades and many switches between targets). We further identify several saccadic signatures that are associated with task difficulty and correct versus incorrect trials for both dimensions. These results suggest distinct encoding algorithms for number and cumulative area extraction, and thereby distinct representations of these dimensions.

## Introduction

Our visual system can quickly, though approximately, represent the number of objects in a scene

through the Approximate Number System (ANS). Visual representations of number obey Weber's law (Cantlon, Platt, & Brannon, 2009; Feigenson, Dehaene, & Spelke, 2004; Halberda & Odic, 2014), are used by human infants and children (Halberda & Feigenson, 2008; Izard, Sann, Spelke, & Streri, 2009; Libertus & Brannon, 2009; Odic, Le Corre, & Halberda, 2015; Odic, Pietroski, Hunter, Lidz, & Halberda, 2013) and many nonhuman animals (Cantlon & Brannon, 2006; Kilian, Yaman, von Fersen, & Güntürkün, 2003; Piffer, Agrillo, & Hyde, 2011; Viswanathan & Nieder, 2013), and have a well-established neural basis in the intraparietal sulcus (Cantlon et al., 2009; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Pinel, Piazza, Le Bihan, & Dehaene, 2004). For example, the ANS allows us to briefly look at the left side of Figure 1 and decide that more of the objects are yellow.

A major open question, however, is how the visual system transforms the retinal input into an appropriate ANS representation—that is, how vision *encodes* approximate number from the visual scene. In other words, which visual features and computations allow ANS representations to form?

The encoding of number has been a highly controversial issue amongst vision scientists, with theories falling into one of two camps: theories suggesting dedicated, number-specific encoding mechanism(s) that extract numeric information directly from the input (Anobile, Cicchini, & Burr, in press; Burr & Ross, 2008; Dehaene & Changeux, 1993; Franconeri, Bemis, & Alvarez, 2009; Park, DeWind, Woldorff, & Brannon, 2015; Ross & Burr, 2010; Sengupta, Surampudi, & Melcher, 2014; Stoianov & Zorzi, 2012), and theories suggesting that number is extracted from one or more nonnumber specific visual features, such as density, surface area, brightness, convex hull (the cumulative surface area covering all the objects), etc. (Cantrell & Smith, 2013; Clearfield & Mix, 1999; Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Durgin, 1995,

Citation: Odic, D., & Halberda, J. (2015). Eye movements reveal distinct encoding patterns for number and cumulative surface area in random dot arrays. *Journal of Vision*, 15(15):5, 1–16, doi:10.1167/15.15.5.

doi: 10.1167/15.15.5

Received June 11, 2015; published November 17, 2015

ISSN 1534-7362 © 2015 ARVO

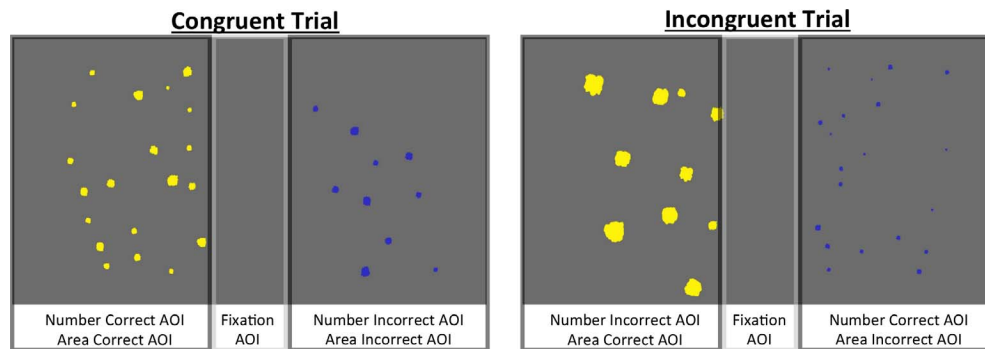


Figure 1. Example of a Congruent and Incongruent trial, alongside the three Area of Interest (AOI) regions. Participants saw identical trials for the Number (more blue or yellow blobs) and Area (more blue or yellow area) Tasks. On Congruent trials the answer by number and cumulative area are the same, whereas on Incongruent trials they differ (the numerically greater side has less cumulative area).

2008; Gebuis & Reynvoet, 2012; Szucs, Nobes, Devine, Gabriel, & Gebuis, 2013; Tibber, Greenwood, & Dakin, 2012). Given that ANS representations must emerge from the extraction of *some* visual feature, a different way of casting the debate is whether number information is or is not a primary visual feature.

In the dedicated encoding theories, the visual system is hypothesized to directly encode number. For example, under several popular models, the visual system extracts and represents the position of each relevant, segmented object (e.g., each blue dot) on a topographic map (Dehaene & Changeux, 1993; Francorneri et al., 2009; Stoianov & Zorzi, 2012). Following a normalization step, the total pooled activity on the topographic map will perfectly correlate with the number of objects: The more objects on the retina, the more positional activity on the topographic map, and the greater the representation of number. Critically, to guarantee that each object contributes the same activity to the pooling computation (e.g., that one big object isn't counted as two small ones), the topographic map must also remove or ignore all nonpositional information from the objects, including their size, shape, brightness, etc. Hence, whereas visual features such as cumulative area and convex hull might correlate with number, the encoding procedure actively suppresses or ignores these nonpositional features. Computationally, the abstraction of nonpositional features by the topographic map can be accomplished either by a prebuilt mechanism (Dehaene & Changeux, 1993), or through learning (Sengupta et al., 2014; Stoianov & Zorzi, 2012). Evidence for object topographic maps has been found, for example, in the LIP area in monkeys (Roitman, Brannon, & Platt, 2007; Stoianov & Zorzi, 2012).

Further evidence for the dedicated encoding theories comes from adaptation studies, where prolonged exposure to a particular stimulus changes the visual system's subsequent response to a novel stimulus (e.g.,

as in color afterimages). If the visual system has a dedicated number encoding mechanism, then one should be able to adapt number-specific neurons. Durgin (1995) was the first to show that prolonged exposure to a high-density display of objects affects their perceived number: High-density adaptation makes the subsequent display appear smaller in number. Subsequent work by Burr and Ross (2008; Ross & Burr, 2010, 2012; but also see Durgin, 2008) showed that—even when density is controlled for—exposure to a stimulus with a large number of objects subsequently makes patches with an identical number of objects appear significantly smaller (e.g., one might misperceive 100 dots as 80). Together, the object topographic map models and the adaptation studies have both been used to support the idea that the visual system encodes number via a dedicated mechanism.

The dedicated encoding models have been criticized, however, and more recently a series of nonnumeric and nonobject specific models have been proposed in their place (Dakin et al., 2011; Gebuis & Reynvoet, 2012; Morgan, Raphael, Tibber, & Dakin, 2014; Szucs et al., 2013). In these models, the visual system is hypothesized to be incapable of representing number directly, and must instead infer the number of objects through nonnumeric visual features such as cumulative area, density, and brightness. These nonnumeric visual features do not depend on object representations and could not be used to make topographic maps, but instead rely on global features such as texture and low spatial-frequency information. As a result, these theories predict that our perception of number will critically depend on the properties of these low-level nonnumeric features, and we should find that number is highly influenced by the differences in, e.g., cumulative surface area. For example, under the model of Dakin and colleagues (2011), number is encoded by combining low spatial-frequency information about density and normalizing it by the convex hull; hence,

the more dense the display and the smaller the objects, the higher the perceived number (i.e., number representations are encoded entirely by proxy through density and area). Number adaptation results could then be explained solely by adaptation of density and/or area neurons, rather than by adaptation of number-specific encoding neurons (Dakin et al., 2011; Durgin, 1995, 2008; but see also Ross & Burr, 2012).

Though perhaps not as initially intuitive, models of visual number encoding through nonnumeric dimensions have both theoretical and empirical support. For example, given that number strongly correlates with nonnumeric features such as size and area (e.g., the more apples are in a bag, the larger the bag is), a dedicated number encoding computation seems unnecessary (Cantrell & Smith, 2013; Gebuis & Reynvoet, 2012). A dedicated system for encoding number also seems neurally implausible. Dakin and colleagues (2011), for example, suggest that number must either be encoded by a yet unidentified item-by-item counting computation, or can otherwise be efficiently estimated through other features, such as combining density and area information, a feat easily accomplished by the primary visual cortex (Dakin et al., 2011; Durgin, 2008; Tibber et al., 2012). Thus, from the perspective of both usability and plausibility, inferring number from dimensions such as area seems theoretically preferable to a dedicated encoding computation.

Several empirical findings also support the view that number is encoded through nonnumeric features. Number discrimination can indeed be affected by manipulating nonnumeric visual features, and especially object size, cumulative area, and convex hull. Hurewitz and colleagues (2006), for example, gave adult observers a number task in which the cumulative area answer was either congruent with number (i.e., the more numerous set of dots was also bigger in area), or was incongruent with number (i.e., the more numerous set of dots was smaller in area). Number discrimination was significantly impacted by congruency, with better performance on congruent trials. Similarly, Durgin (1995) observed that number representations are related to density: The denser the display, the higher the estimated number. Other work utilizing congruent/incongruent designs has shown that visual features such as convex hull, object diameter, and object size all affect number discrimination and estimation (Gebuis & Reynvoet, 2012; Szucs et al., 2013).

But, despite all the work to date, evidence for the dedicated versus the nonnumeric feature encoding theories has been equivocal. For example, whereas the number-specific adaptation findings have been criticized for tapping into features like density (Dakin et al., 2011; Durgin, 2008; Tibber et al., 2012), other work has suggested that number adaptation occurs even when density is controlled for (Anobile et al., in press; Ross &

Burr, 2010, 2012) or only when displays contain well over twenty objects (Anobile, Cicchini, & Burr, 2014). Additionally, altering nonnumeric visual features such as density may also incidentally impact object maps, as denser collections include objects that are more crowded and harder to extract (Allik & Tuulmets, 1991; Anobile et al., 2014, in press; Ross & Burr, 2012). Recent EEG evidence also suggests that the early visual cortex is more sensitive to changes in number than other visual features (Park et al., 2015). And, whereas some findings report a difference between congruent and incongruent trials (Gebuis & Reynvoet, 2012; Gilmore et al., 2013; Szucs et al., 2013), many have failed to find any differences, especially when the ratio of the nonnumeric feature is matched to that of number (Barth, 2008; Libertus, Odic, & Halberda, 2012; Odic, Hock, & Halberda, 2014; Odic, Libertus, Feigenson, & Halberda, 2013; Piazza et al., 2004). Finally, the congruency effects may not be indicative of number encoding, but instead may reflect the observer's failure to use the intended dimension (i.e., number) when other dimensions are salient and/or easier for the trial at hand (Gilmore et al., 2013).

In the experiment reported here, we test the similarities and differences in number and cumulative area encoding by monitoring eye-movements while observers perform either a number discrimination task ("Are more of the blobs blue or yellow") or a cumulative area discrimination task ("Is more of the blob blue or yellow"). By monitoring eye-movements, we can determine whether the information observers seek in the display is identical or different when discriminating number and when discriminating cumulative surface area. For example, if participants infer number through a nonnumeric dimension such as area, eye-movement patterns during the number task should be similar to the patterns observed during the area task. Critically, we give observers identical displays for both of these tasks, allowing us to directly compare their saccadic movements when only the judged dimension differs.

What saccadic signatures might we expect if the number and cumulative area encoding computations are distinct versus if they are similar? Because no previous study has examined performance on number and area discrimination tasks while concurrently tracking eye-movements, our predictions of the relevant signatures are entirely informed by the models reviewed above.

In the case of area encoding, there is a high degree of agreement between the two theories: Surface area is most likely encoded by pooling texture and low spatial-frequency information using global, distributed attention (Alvarez, 2011; Cant & Xu, 2012; Chong & Treisman, 2003, 2005; Corbett & Melcher, 2013; Dakin et al., 2011; Haberman & Whitney, 2012).<sup>1</sup> Because the

processing of low spatial-frequency information is equally good in peripheral and foveal vision (Strasburger, Rentschler, & Jüttner, 2011), and because saccades are rare with distributed attention (Kowler, Anderson, Doshier, & Blaser, 1995; McPeck, Maljkovic, & Nakayama, 1999; Rizzolatti, Riggio, & Sheliga, 1994; Weber, Schwarz, Kneifel, Treyer, & Buck, 2000), area encoding should produce very few saccades, and hence, very long gazes.

If number is encoded through surface area (either directly, as predicted by Gebuis & Reynvoet, 2012, or with area as a normalizing function for density, as predicted by Dakin et al., 2011), then the encoding of number should similarly depend on pooling texture and low spatial-frequency information using distributed attention. As a result, if number is inferred through cumulative area encoding, eye-movements should be similar for both tasks. First, participants should spend more time looking to the side with higher cumulative area (as that side should also be judged to have more objects). Second, participants should show few saccades and long gazes, indicative that they are encoding number through distributed attention and the pooling of low-level spatial frequency. In summary—if number is inferred by using evidence from cumulative surface area—eye movements should be near-identical on the number and an area tasks.

The dedicated encoding theory, however, predicts a different pattern of eye-movements for number compared to cumulative surface area. Since dedicated encoding models critically depend on identifying exact object positions, number encoding will suffer in conditions where objects positions cannot be easily identified or where multiple objects are close enough that they cannot be reliably bound into proto-objects or segmented into object representations (He, Cavanagh, & Intriligator, 1996; Rensink, O'Regan, & Clark, 1997; Walther & Koch, 2006). Such proto-object or object representations may be disrupted either when the objects are densely packed together (Anobile et al., 2014; Ross & Burr, 2012), or when they are in peripheral vision, where visual crowding, eccentricity, and distributed attention will interfere with individuation (Anobile et al., 2014; He et al., 1996; Rosenholtz, Huang, Raj, Balas, & Ilie, 2012; Ross & Burr, 2012). As a result, number encoding theories that depend on representing object positions (e.g., on a topographic map) predict that participants will make many quick saccades, including to peripheral areas of the visual display, in an effort to separate crowded objects and correctly represent their positions on the topographic map. As a result, dedicated number encoding should yield many saccades, and hence short gaze durations, especially when compared to cumulative area encoding, and should also result in greater looking to the array

that is greater in number—irrespective of any differences in total area.

## Experiment

### Methods

#### Participants

12 adult observers participated for course credit. Three additional participants were removed because the eye-tracker could not reliably track their eyes (i.e., we observed fewer than 10% good samples).

#### Materials

Participants were tested on a Tobii TX300 running at 300 Hz. The eye-tracker was mounted on a 22" widescreen monitor. The stimuli were presented with Tobii Studio 3.2 running on Windows 7.

Each participant was given both a Number Task and an Area Task, in counterbalanced order. In order to make sure that none of our eye-tracking results were affected by difficulty or response time, we created stimuli that (in pilot testing) showed identical average discrimination performance and average RT for the two tasks. The stimulus displays used in the two tasks were identical, and consisted of many blue and yellow blob shapes (see Figure 1); all of the blue blobs were presented in an invisible box centered approximately 200 pixels to the right of fixation, and all of the yellow blobs were presented in an invisible box centered approximately 200 pixels to the left of fixation.

To modulate difficulty, we manipulated the ratio of blue to yellow blobs, and the ratio of cumulative blue to yellow area (i.e., number of pixels). Ratio was defined as the larger quantity (number, pixels) divided by the smaller quantity. For example, the stimulus on the left side of Figure 1 has 20 yellow versus 10 blue objects (a ratio of 2.0), and 1700 yellow pixels versus 850 blue pixels (a ratio of 2.0). In total, we used five ratios: 2.0, 1.5, 1.2, 1.14, and 1.12, with each trial's number ratio matching the trial's area ratio.

Half the trials were Congruent (number and cumulative surface area agreed on the answer; e.g., Figure 1, left), and half were Incongruent (number and cumulative surface area disagreed on the answer; e.g., Figure 1, right). Congruent trials also matched average object size for the two sets.

#### Procedure

Participants were calibrated with an 8-dot calibration before the start of each task. Subsequently, each participant did both the Number and the Area Task,

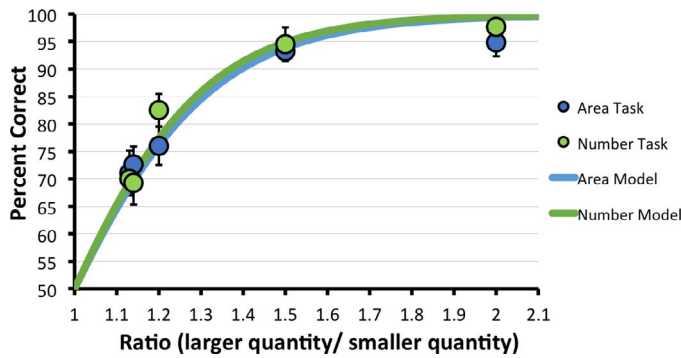


Figure 2. Average performance on the Number (green) and Area (blue) Tasks. Lines represent the standard psychophysical model with each dimension's Weber fraction as the only free parameter. Error bars are *SEM*.

with order counterbalanced across subjects. During the Area Task, participants were asked: “Is more of the blob blue or yellow,” and to press the “F” key for “More of the blob is yellow,” and the “J” key for “More of the blob is blue.” In other words, participants had to verify whether the cumulative area (i.e., total area of pixels) was higher for the blue or yellow side, independent of number. In the Number Task, they were asked: “Are more of the blobs blue or yellow,” and to press the “F” key for “More of the blobs are yellow,” and the “J” key for “More of the blobs are blue.” In other words, participants had to verify whether the total number of items on the screen was higher for the blue or yellow side, independent of cumulative area. The experimenter stayed with the participant for four practice trials in order to make sure that the participant understood which dimension to attend to on each Task.

Each trial was presented for 2000 milliseconds (ms), and participants were able to respond at any time during that period. The stimulus stayed up after the response, allowing eye tracking to continue for the full 2000 ms period. However, all of our reported results are truncated to the participant's response time on the particular trial. Each ratio was presented in 32 distinct trials, yielding 160 total trials per Task.

### Eye-tracking analysis

All eye-tracking variables were extracted from the data by examining fixation points along the horizontal and vertical axes; fixation points were determined with thresholds on both velocity and position. As shown in Figure 1, we created three main areas of interest (AOIs) depending on the task and correct answer: the correct side, the incorrect side, and the fixation (which took up the central 20% of the screen). Any samples that were unreliable (e.g., because the participant blinked or looked off-screen) were removed from the data analysis (9% of total data).

The raw data files were analyzed with custom-made MATLAB scripts. The primary variables of interest included the onset of the first saccade; the proportion of time spent looking at each AOI (e.g., correct vs. incorrect); the location and the duration of the first, last, and longest fixation; and the number of switches between the AOIs. Although pupil size was also measured, it did not correlate with any measure.

## Results

The results are presented in three sections. First, we report the behavioral effects of Task on average performance and RT (ignoring eye-tracking data). Second, because no previous study has measured eye movements during number or cumulative area discrimination, we report a number of signatures that were identical for the two Tasks, including effects of ratio. Finally, we combine the eye-tracking data with Task to identify any saccadic signatures that might differentiate number from area encoding.

For the eye-tracking analysis, we focused on two broad categories of saccadic signatures: those identifying *where* participants looked, and those identifying *how* participants looked. For examining saccade locations, we focused on the percentage of time spent at each AOI (Correct, Incorrect, and Fixation) and the AOI associated with the longest gaze. If number is encoded straightforwardly through cumulative area, we would expect that participants spend most of their fixations looking at the AOI with more cumulative area, even when discriminating number (e.g., because the side with greater area would be experienced as greater in number). For examining the nature of saccades, we examined the total number of fixations, the percent of fixations spent switching between any two AOIs, saccadic onsets (time until the first saccade is made), and the duration of the longest gaze. As we detail below, area encoding predicts few fixations and switches, but long onsets and gaze durations. In contrast, a dedicated number encoding algorithm—that is independent from area—predicts many fixations and switches, but comparatively short onsets and gaze durations. Note that all signatures confounded with response times were normalized as percentage of total fixations, avoiding this issue (though all our results remain identical whether RT is controlled for or not).

### Accuracy and RT analyses (behavior)

Performance on the Number and Area tasks across ratio is shown in Figure 2. A  $2$  (Order: Number-First, Area-First)  $\times 2$  (Task: Number, Area)  $\times 5$  (Ratio: 2.0, 1.5, 1.2, 1.14, and 1.12) mixed-measures ANOVA over percent correct showed no main effects or interactions

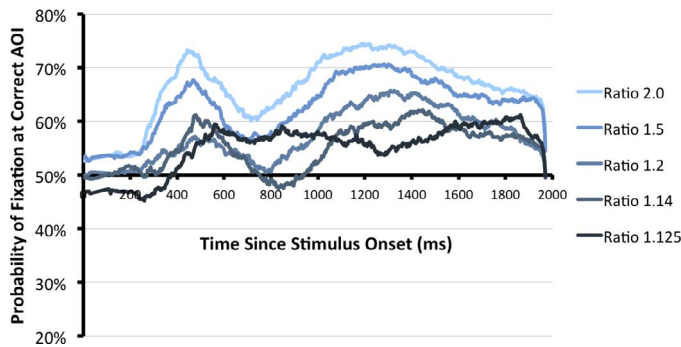


Figure 3. The probability of a fixation being on the correct AOI as a function of Ratio, collapsed across both the Number and Area Task. Thus, higher values on the y axis indicate more fixations on the side with more number or cumulative area. Note that the data plotted here are not cropped at moment of response (as the RT differed for each participant), whereas the analyses are.

with Order (all  $p$ s > 0.15); we removed Order from future analyses. As expected from our pilot testing, there was no main effect of Task,  $F(1, 11) = 2.08$ ,  $p = 0.18$ , with equivalent performance on the Area (Mean = 0.80;  $SE = 0.03$ ) and Number Tasks (Mean = 0.83;  $SE = 0.02$ ). As expected via Weber's law, there was a significant effect of Ratio,  $F(4, 44) = 55.49$ ,  $p < 0.01$ , allowing us to estimate each participant's Weber fraction (Halberda & Feigenson, 2008; Halberda & Odic, 2014; Piazza et al., 2004). The average Weber fraction for the Area Task was 0.14 ( $SE = 0.02$ ), and for the Number Task was 0.13 ( $SE = 0.01$ ).

We found no difference between the Number and Area Tasks' response times. A 2 (Task: Number, Area)  $\times$  5 (Ratio: 2.0, 1.5, 1.2, 1.14) mixed-measures ANOVA over RTs showed an expected main effect of Ratio,  $F(4, 48) = 47.69$ ,  $p < 0.001$ , with RTs decreasing as a function of Ratio. There was no main effect of Task,  $F(1, 12) < 1$ , nor a Task  $\times$  Ratio interaction,  $F(4, 48) = 2.13$ ,  $p > 0.10$ .

We found that the Number task was unaffected by cumulative area (Congruent vs. Incongruent), while performance in the Area task was superior on Incongruent trials (i.e., where number and area conflicted across the array). A 2 (Task: Number, Area)  $\times$  2 (Congruency: Congruent, Incongruent) mixed-measures ANOVA showed a significant Congruency  $\times$  Task interaction,  $F(1, 12) = 8.56$ ,  $p < 0.02$ : The Number Task was not impacted by Congruency (Congruent: 88.5%,  $SE = 1.2$ ; Incongruent: 88.4%,  $SE = 1.72$ ) but the Area Task Congruent performance was worse than the Area Task Incongruent performance (Congruent: 81.9%,  $SE = 1.7$ ; Incongruent: 91.0%,  $SE = 2.01$ ). The lack of a congruency effect in the Number Task - i.e., identical performance when changing cumulative area—is consistent with a dedicated encoding mechanism and the results of Barth (2008), Odic

and colleagues (2013), and others, but inconsistent with the results of Gebuis and Reynvoet (2012), Hurewitz and colleagues (2006), or Szucs and colleagues (2013). The inverse effect of Congruency on the Area Task (i.e., better Incongruent performance) is unexpected by any theory and may stem from Incongruent trials having larger objects on the correct side, allowing for more efficient low spatial-frequency extraction compared to when objects were smaller. Alternatively, participants may have used both cumulative area and average object size as a cue on the Incongruent trials, but could use only cumulative area on the Congruent trials, where object size average was controlled.

### Effects of ratio and accuracy on saccadic signatures of both tasks (eye-tracking)

Because no previous experiment has reported eye-tracking patterns during approximate number or cumulative area discrimination, we report how a number of factors, including difficulty and accuracy, influence saccades on both tasks. Subsequently, we focus on saccadic signatures that differentiate number from area encoding.

Task difficulty—that is, ratio—affected both where participants looked and how participants looked at the displays: Easier ratios were associated with more looks to the correct AOI and many fewer switches between AOIs. As can be observed from Figures 3 and 4, participants spent most of their time looking at the correct AOI, but this effect decreased with harder ratios,  $F(4, 48) = 3.96$ ,  $p < 0.01$ . Furthermore, the probability of the very first fixation being to the correct AOI also increased with Ratio,  $F(4, 48) = 6.87$ ,  $p < 0.01$ , as did the probability of the longest fixation being on the correct AOI increased with Ratio,  $F(4, 48) = 9.851$ ,  $p < 0.001$ . Participants spent many more fixations switching between AOIs with harder ratios compared to easier ones, most likely in an effort to identify the correct AOI (Figure 4, upper right panel;  $F(4, 48) = 21.65$ ,  $p < 0.01$ : At the hardest ratio (1.125), participants spent approximately 33% ( $SE = 1.9\%$ ) of their fixations switching between the AOIs, whereas at the easiest ratio (2.00) they spent about 23% ( $SE = 2.1\%$ ) of their fixations switching between the AOIs. Note also that the significant effect of RT over Ratio could itself be the product of a higher number of fixations and switches before participants decide their answer; indeed, RT and the total number of fixations significantly correlated;  $r(23) = 0.60$ ,  $p < 0.001$ .

If participants generally spend more time looking at the correct AOI, do they also spend more time looking at the incorrect AOI when they answer incorrectly? Figure 5 shows the probability of fixation being to the correct AOI as a function of time and whether the participant's answer was correct or incorrect. The

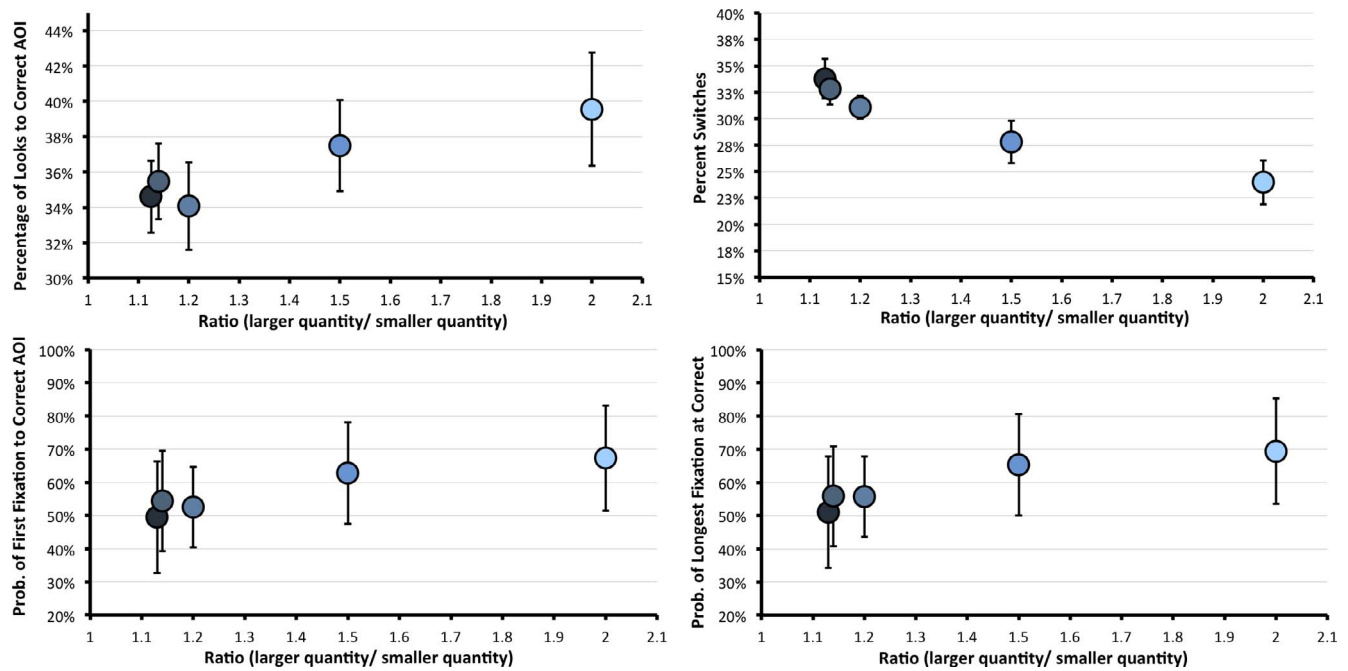


Figure 4. The effects of Ratio, collapsed across Number and Area Tasks, on four saccadic signatures: (Upper Left) the average percentage of fixations to the correct side, (Upper Right) the number of switches between the two target sides, (Lower Left) the probability of the very first fixation being to the correct side, and (Lower Right) the probability of the longest fixation being on the correct side. All error bars are SEM.

pattern is clear: When participants answer incorrectly, they are also more likely to be looking at the incorrect AOI. Similarly, the percentage of time participants spend looking at each of the three AOIs is distinct for correct versus incorrect trials (Figure 5): On correct trials, participants spend 37.6% ( $SE = 0.03$ ) of time in the correct AOI, 18.8% ( $SE = 0.03$ ) of time in the incorrect AOI, and 43% ( $SE = 0.03$ ) in the fixation AOI; on incorrect trials, they spend 26.9% ( $SE = 0.02$ ) in the correct AOI, 36.1% ( $SE = 0.02$ ) in the incorrect AOI, and 36.9% ( $SE = 0.02$ ) in the fixation AOI. The difference between percentage of time spent in the correct AOI as a function of being correct or incorrect

was significant,  $t(12) = -3.24, p < 0.01$ . Additionally, all of the saccadic signatures discussed above are affected by incorrect responses: The probability of the first fixation being to the correct side was significantly lower on incorrect trials, 46.5% versus 58.7%,  $t(12) = -2.25, p < 0.05$ ; the probability of the longest fixation being on the correct side was significantly lower for incorrect trials, 34.8% versus 62.9%,  $t(12) = -4.26; p < 0.001$ , and the proportion of switches between AOIs was significantly higher for incorrect trials, 33.9% versus 27.3%,  $t(12) = -3.40, p < 0.001$ .

Overall, the results here suggest that how participants ultimately answer is strongly related to where

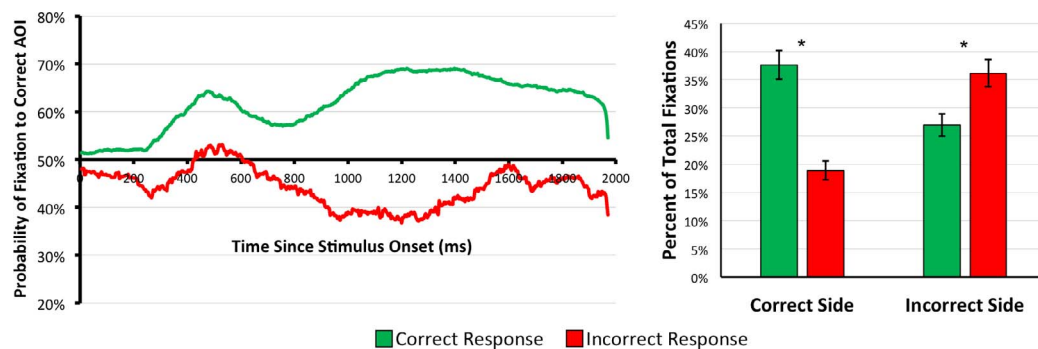


Figure 5. The probability of fixation being to the correct side, collapsed across Number and Area Tasks (left), and a bar graph illustrating the percent of fixations to the correct versus incorrect side as a function of whether the participant answered correctly or incorrectly, collapsed across the Number and Area Tasks. Bars are SEM. Stars indicate  $p < 0.05$ . As can be seen in the graph, incorrect answers are associated with more looks to the wrong side.

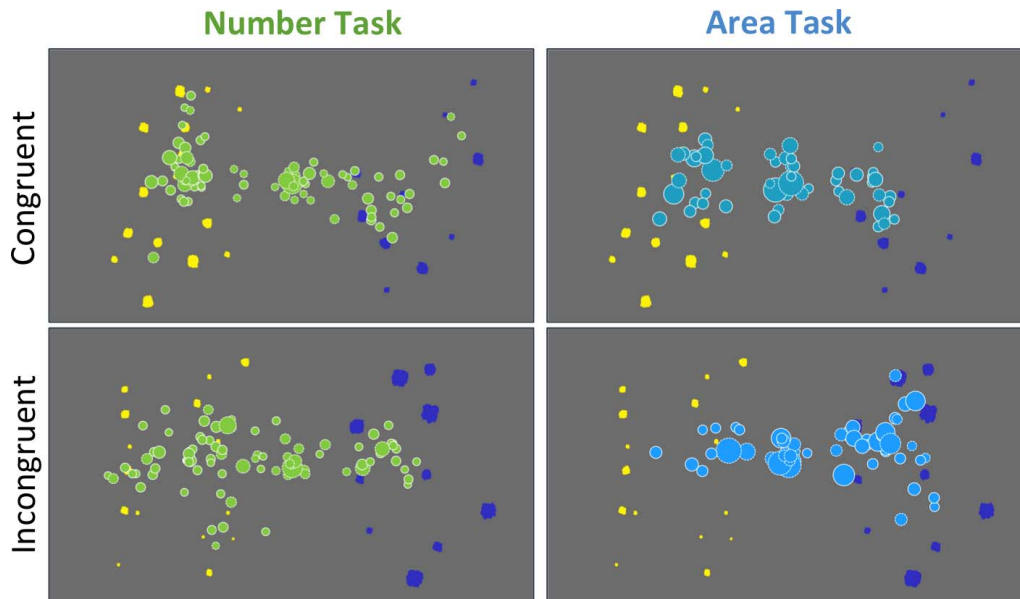


Figure 6. Four cumulative fixation plots, collapsed across all participants, displaying each look made during the task (i.e., during either the Area or Number question block) for two example arrays. Each image is scaled down from the actual display while preserving the aspect ratio, with the fixation point being in the middle of each display. Each dot indicates a fixation, with the size of the dot indicating the duration of gaze. On congruent trials, the side with more objects also has more cumulative area. On incongruent trials, the side with fewer objects has more cumulative area. As can be seen from these example images, Task had a significant effect on saccadic signatures: Participants differ in where they look and how they look, with number showing many more short saccades compared to area.

they look: The more time they spend looking at a particular AOI, the more likely they are to select it as the answer to the “more” question. Furthermore, on harder trials (i.e., lower ratios) participants switch significantly more between the AOIs, presumably reflecting uncertainty and an attempt to identify the correct side. On incorrect trials, participants’ saccadic signatures are strikingly similar to those on correct trials, but in the opposite direction. This suggests that participants are likely not blindly guessing when ratios are hard, but that they settle on the incorrect side and judge it to be higher in the relevant quantity.<sup>2</sup>

#### **Effect of task on eye-tracking signatures (eye-tracking)**

Finally, we turn to the most important set of results: Are there any saccadic signatures that can differentiate number from area encoding?

Figure 6 illustrates an example for four trials—one congruent and one incongruent, one from the Number Task and one from the Area Task—of the typical patterns of saccades and gazes seen for each Task. Each dot represents a saccadic position (collapsed across all twelve participants), while the size of the dot indicates fixation duration at that location. As elaborated in detail below, these images illustrate the typical pattern observed across all trials: The encoding of number and area can be distinguished both by *where* participants

look (e.g., the primary AOI attended during Incongruent trials is different for the two tasks) and *how* participants look (i.e., the number task shows significantly more quick saccades (smaller disks in Figure 6), including to the edge of the cloud of objects).

If participants are inferring number from cumulative area, we would expect that participants spend most of their time looking to the side with more cumulative area, even when they are judging number (as this side should appear to have more objects, as well). But—contrary to these predictions and consistent with dedicated encoding models—number encoding resulted in a greater probability of looks to the side that was greater in number, whereas area processing resulted in a greater probability of looks to the side that was greater in cumulative area (Figure 7). Hence, one saccadic signature that clearly differentiates number from area discrimination is location: Participants spend most of their time looking at the side with more number in the Number Task, and more area in the Area Task. As discussed above, this result makes it unlikely that participants are encoding or inferring number *through* cumulative area, as this might predict that they should be spending most of their time looking to the side with the stronger area signal.

We also found that number encoding changed *how* participants looked at the displays. Consistent with the dedicated encoding account, the Number Task was associated with many more saccades and shorter gazes,



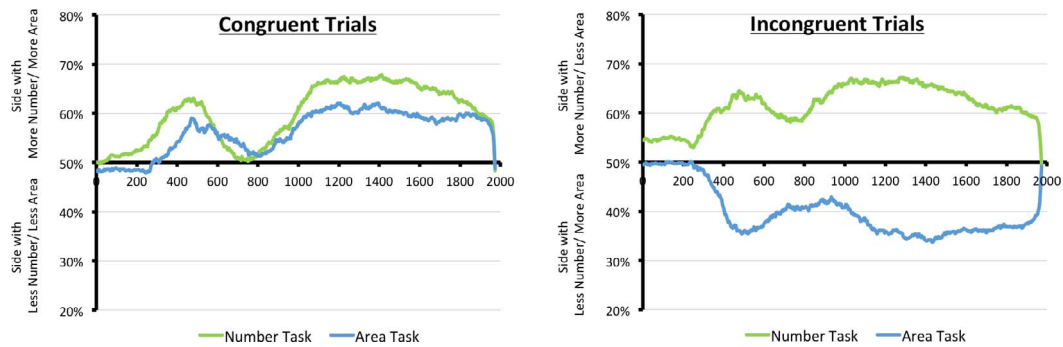


Figure 7. Probability of looking to the side with more number or more area as a function of Task and Congruency, ignoring the center screen region (chance = 50%). In the Congruent trials, participants’ looking in Area and Number tasks is highly similar, with both tasks resulting in looks that gravitate towards the side greater in number/area. In the Incongruent condition, the opposite result is seen whereby looking on the Area and Number tasks strongly diverges with participants looking to the side that is correct under their currently relevant dimension. Note, if number were represented via area, this is an unexpected result. However, if number has its own representation and extraction algorithm that is independent of area, then this is a predicted result.

while the Area Task was associated with many fewer saccades—including much more time spent at fixation—and much longer gazes. This difference in saccadic signatures between the Number and Area Tasks is empirically demonstrated through three signatures: central fixation time, gaze duration, and number of switches (Figure 8).

First, we found that the participants spent much more time at the fixation AOI in the Area Task, consistent with the prediction that area encoding depends on distributed attention: As shown in Figure 8, the percentage of time participants spend in the fixation AOI is significantly higher for the Area Task (46%,  $SE = 0.05$ ), than for the Number Task (39%,  $SE = 0.05$ ) as measured by paired-samples  $t$  test,  $t(12) = 2.82, p <$

0.02; consistent with number encoding yielding more saccades around the screen. This effect holds for both Congruent and Incongruent trials: A 2 (Task: Number, Area)  $\times$  2 (Congruency: Congruent, Incongruent) repeated-measures ANOVA over percent of time spent in the fixation AOI showed a main effect of Task,  $F(1, 12) = 7.77, p < 0.05$ , but no interaction,  $F(1, 12) = 3.10, p = 0.10$ . Relatedly, there was a marginally significant effect of the onset to first saccade, which was delayed in the Area Task, 380 ms,  $SE = 38.0$ , compared to the Number Task, 345.20 ms,  $SE = 31.5, F(1, 12) = 4.20, p = 0.063$ .

Second, participants showed much longer gazes in the Area Task compared to the Number Task, consistent with the dedicated encoding theory: As

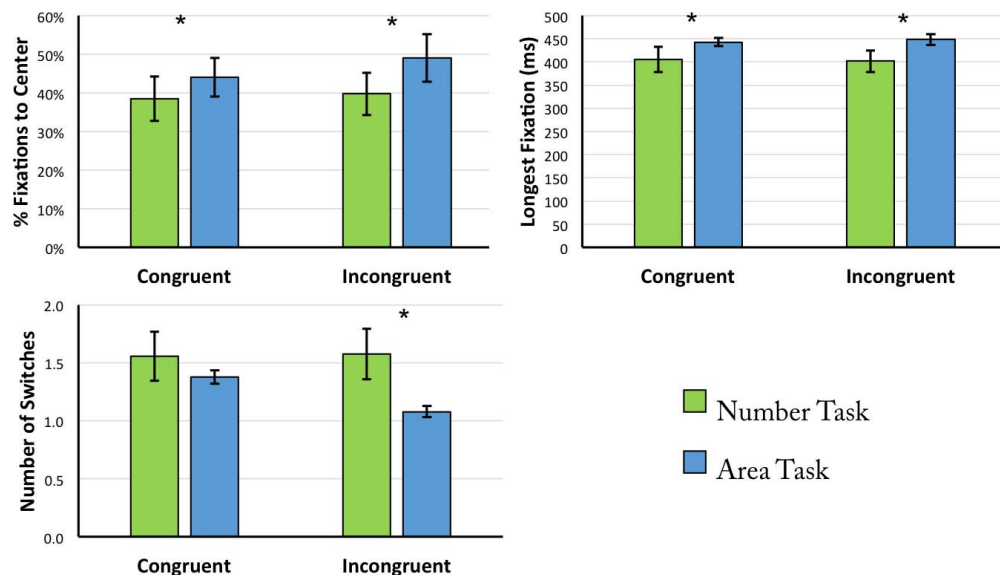


Figure 8. The effects of Task on three saccadic signatures: (Upper Left) average percent of time spent in fixation, (Upper Right) the duration of the longest fixation truncated to RT, (Lower Left) the number of switches between the two target sides. All error bars are SEM. Stars indicate  $p < 0.05$ . The Area Task is associated with longer fixations and fewer switches.

shown in Figure 8, the duration of the longest fixation (after the first saccade) was significantly longer in the Area Task (444.84 ms,  $SE = 32.35$ , than in the Number Task, 403.46 ms,  $SE = 25.20$ ,  $t(12) = 2.22$ ,  $p < 0.05$ . Once again, this effect did not show an interaction with Congruency: We observed a main effect of Task,  $F(1, 12) = 5.01$ ,  $p < 0.05$ , but no main effect of Congruency nor an interaction (both  $F_s < 1.0$ ).

Finally, participants showed many more switches between the AOIs in the Number Task compared to the Area Task: As shown in Figure 8, the number of switches between the two sides was significantly higher in the Number Task, 1.56;  $SE = 0.21$ , compared to the Area Task, 1.22;  $SE = 0.17$   $t(12) = 2.31$ ,  $p < 0.05$ . This effect, however, was primarily driven by the difference on Incongruent trials: There was no main effect of Task,  $F(1, 12) = 2.79$ ,  $p = 0.12$ , but there was a significant interaction with Congruency,  $F(1, 12) = 20.88$ ;  $p < 0.01$ . The absence of this effect on Congruent trials is likely driven by the Area Task Congruent trials being significantly harder (recall that harder trials generally show more switches). Hence, the effect of switches on Incongruent trials (which did not significantly differ in accuracy or RT for the Number vs. Area Tasks) is further validation of the differences in switches between number and area encoding.

Our results can be synthesized into two major findings. First, participants in the Number Task—unlike the Area Task—spend most of their time looking to the side with more objects. Second, participants in the Number Task show—compared to the Area Task—many more saccades and switches, including fewer looks to the fixation AOI, consistent with number encoding depending on focal attention and area encoding depending on distributed attention. These results dissociate the encoding of number and area and are most consistent with the dedicated encoding models.

A possible criticism of our interpretation of these results could be that number encoding occurs through a combination of cumulative area and other visual features, such as brightness or low spatial-frequency (Gebuis & Reynvoet, 2012). For example, perhaps participants quickly encode cumulative area, and then subsequently encode other dimensions, thus increasing switches and/or other saccadic signatures away from pure cumulative area encoding. This interpretation is inconsistent, however, with two aspects of our results. First, participants performing incongruent Number Task trials spent almost all of their time (including their first saccade) on the side with more objects, contrary to the critical signatures of area encoding in our stimuli (see Figure 7). Second, if area information is combined with other features during number encoding, the Number Task saccadic signatures should be additive with those of the Area Task (i.e.,

number encoding will include all signatures of area encoding, plus additional ones). But, this is the exact opposite of what we find: Several saccadic signatures, including percent of fixations in the central AOI and the duration of longest fixation, are *lower* in the Number Task than in the Area Task, making additivity impossible. And, given that the two tasks had identical accuracy and RT, the saccadic signatures observed for area do not appear to include idling time that would allow for additivity. Finally, given that the stimuli were identical, if number is encoded via surface area, then surface area performance would place an upper bound on number performance with decrements occurring with each additional factor required for a number judgment. In conflict with this prediction, number performance was just as good as area performance in our task, and with a distinct set of eye-movement signatures supporting successful number discrimination.

Together, these results suggest that number and area encoding not only promote differences in *where* participants look (i.e., to the side with more number vs. to the side with more cumulative area), but also in *how* they look (i.e., with number showing many more saccades, and much quicker looks). As discussed below, these results are most consistent with the dedicated encoding theory for number cognition.

## Discussion

Theories of visual number encoding fall into one of two categories: Dedicated encoding theories claim that the visual system can directly extract number from the retinal input (e.g., via positions of objects on a topographic map), whereas nonnumerical theories claim that the visual system infers number from a range of nonnumeric and nonobject visual features (e.g., surface area). In this experiment, we tracked eye movements of participants who performed both a number and a cumulative surface area discrimination task over identical stimuli, and sought to identify whether the visual information explored in the displays was identical for number and area (as predicted by some nonnumerical theories) or if it was different (as predicted by dedicated encoding theories). Our results were more consistent with the dedicated encoding theories.

We found that, even with identical displays, asking participants to discriminate via number produced significant differences in saccadic signatures compared to when they were asked to discriminate via cumulative surface area. First, whether participants discriminated number or cumulative area modulated *where* participants looked: During the number task, participants

looked significantly longer and more often to the side that had more objects, while during the area task they looked significantly longer and more often to the side that had more cumulative surface area. Thus, participants did not appear to encode number by looking to the side with the stronger cumulative area signal. Task also modulated *how* participants looked at the displays: The number task was associated with significantly more and quicker looks, and more switches between the targets; the area task was associated with significantly fewer, long saccades, and fewer switches between targets (see Figures 6 and 8). These results cannot be explained by differences in congruency, accuracy, or RT, and are best explained by a difference in the underlying encoding procedures: Number encoding depends on more saccades in an effort to gather object-related information and attend/discriminate among crowded items, whereas area encoding depends more on global processing of nonnumeric features and distributed attention.

Placing our data in a broader context, we emphasize that our results do not suggest an exact algorithm used by the visual system to extract number. Because the world is dynamic, there is rarely a single extraction algorithm that will fulfill the needs of cognition across all contexts. It is true that, in many real-world cases, area is a good proxy for number. And, if the extraction of area information required fewer resources than the extraction of number, then it would be a rational choice to use it. But area and number can be made to conflict with each other—as was done in the Incongruent trials of the present experiment. In such cases, vision must either stick with a failing algorithm (e.g., cumulative area), or dynamically adjust to gain better purchase on the dimension of interest (e.g., what is the number of blue dots?). Our results are consistent with either the possibility that humans are able to make rapid adjustments in the algorithms they use to estimate number, or with the possibility that a single number-specific visual algorithm normalizes appropriately across the variety of contexts to arrive at an estimate of visual number that is robust to changes in cumulative area. Furthermore, our study only investigated the encoding algorithms of *visual* number; as number representations are well known to be amodal (Izard et al., 2009; Jordan & Brannon, 2006; Meck & Church, 1983), an entirely different set of algorithms will likely be required to describe the auditory encoding of number.

Beyond our findings on signatures that differentiate number from cumulative area encoding, we also found saccadic signatures that were shared between them and depended on ratio and whether participants answered correctly or incorrectly. First, the proportion of time spent looking to the correct side decreased with harder ratios, while the number of switches

between AOIs increased. Second, when participants answered incorrectly on the trial, the saccadic signatures were inverted compared to when they answered correctly: They looked longer to the incorrect side, and spent even more time switching between AOIs. Together, these results indicate that an observer's ultimate response aligns with where they look the most. Additionally, the inverse signatures during incorrect trials suggest that these are unlikely to be pure guesses, and instead may be trials on which the incorrect side truly looked more numerous or larger in area.

The results reported here are broadly consistent with more recent work suggesting independence between estimating number and surface area. Work from our lab has shown that the Weber fraction (i.e., precision) of number and cumulative area representations are independent within individual observers, even across development (Odic, Libertus, et al., 2013). Much like in the results reported here, there was also no effect of area-congruent versus area-incongruent trials on number discrimination at any age. Lourenco, Bonny, Fernandez, and Rao (2012) found that individual differences in number Weber fractions uniquely correlated with arithmetic math problems, whereas individual differences in approximate cumulative surface area uniquely correlated with geometric math problems. Finally, work from Castelli, Glaser, and Butterworth (2006) has shown that unique substrates of the intraparietal sulcus code for number versus surface area. Together, these results suggest that number processing is importantly independent from cumulative area processing, and that cumulative area is not contributing to the encoding of number.

If number and area encoding rely on different computations, why do some studies find differences in visual number discrimination on congruent versus incongruent trials? Given that this difference is not always reliable (e.g., we failed to observe it here), the reported differences may be explained in at least two ways. First, *some* manipulations of surface area or density may actually impair number encoding, even if the encoding procedures are different (e.g., small objects in the periphery may be missed or difficult to represent on object maps; Ross & Burr, 2010). Tracking saccadic signatures may allow us to test this hypothesis in the near future: Given that more difficult number trials are associated with more saccades and switches, one expectation would be that manipulations of area that strongly impact number encoding should also increase the number of saccades and switches, even when the number ratio is kept constant. This would be the opposite of the pattern predicted if participants are instead using surface area as a cue to number (in which case, the number of saccades and switches should decrease as cumulative area becomes

easier to encode). As we did not find a difference between congruent and incongruent number trials in this experiment, this remains an avenue for future research.

The second possibility for the occasional congruency differences may be that nonnumeric dimensions are sometimes useful cues to number: As discussed above, because number and total surface area are highly correlated in the world, it would be rational for the observer to use surface area if computing it is easier and quicker than computing number (e.g., if the ratio for cumulative area is easier than the one for number). For example, we would expect that individuals with poorer ANS precision might compensate in daily life by relying on surface area. Hence, the congruency difference may not be indicative of an area-based number encoding, but rather a rational substitution choice of one dimension over another in order to perform better on most tasks. In this case, number discrimination saccadic signatures, such as the high number of switches, should decrease as area becomes more salient and easier to use.

An important caveat of the current work is that we focused entirely on how saccadic signatures differentiate number from surface area. Although these dimensions were motivated by the majority of nonnumeric feature theories and models focusing on surface area (either as a cue to number directly, or as a normalizing procedure for density; Dakin et al., 2011; Gebuis & Reynvoet, 2012), it remains an open possibility that number is encoded through visual features that do not involve surface area (e.g., contrast, diameter, etc.). Future work utilizing eye tracking during discrimination performance could be used to test these visual features. Any such work must explain why number encoding demonstrates an increased number of saccades and switches between targets, as an object-based encoding account provides a straightforward explanation to this pattern.

In summary, our work is the first to report eye-tracking data applied to a highly controversial and popular question: How does the visual system encode number? Although the evidence presented here suggests that the algorithms that encode for number are distinct from those that encode for cumulative area, this work cannot—at the moment—specify the exact nature of visual number encoding. But, by adding eye-tracking data into a literature that already combines neuroscience, psychophysics, and cognitive, computational, comparative, and developmental psychology, we hope to provide another way that researchers can come together and understand the nature of number in vision and cognition.

*Keywords:* approximate number system, area perception, eye-tracking

## Acknowledgments

This research was funded by NICHD R01 HD057258 awarded to J. H., and by a NSERC PGS-D awarded to D. O. Thank you to Klaus Libertus and Melissa Libertus for help in data analysis, and to Michelle Hall for useful advice on the manuscript.

Commercial relationships: none.

Corresponding author: Darko Odic.

Email: darko.odic@psych.ubc.ca.

Address: Department of Psychology, University of British Columbia, Vancouver, BC, Canada.

## Footnotes

<sup>1</sup> The use of nonobject features, such as texture and low spatial-frequency, is consistent with findings that estimates of total or average object size are superior to estimates of a single object's size (Ariely, 2001; Chong & Treisman, 2003). The use of distributed attention during area encoding is consistent with decreases in area performance when participants are forced to engage in focal attention (Chong & Treisman, 2005; Joo, Shin, Chong, & Blake, 2009).

<sup>2</sup> An additional analysis examined whether any features of the display led the majority of participants to respond incorrectly on specific trials. However, we found little consistency in participant's responses on specific trials: We identified only two trials that showed below-chance performance when averaged across all participants.

## References

- Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Attention, Perception, & Psychophysics*, *49*(4), 303–314. doi.org/10.3758/BF03205986.
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131.
- Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science*, *25*(1), 265–270. doi.org/10.1177/0956797613501520.
- Anobile, G., Cicchini, M., & Burr, D. C. (in press). Number as a primary perceptual attribute: A review. *Perception*, in press.
- Ariely, D. (2001). Seeing sets: Representation by

- statistical properties. *Psychological Science*, *12*(2), 157–162, doi.org/10.1111/1467-9280.00327.
- Barth, H. (2008). Judgments of discrete and continuous quantity: An illusory Stroop effect. *Cognition*, *109*(2), 251–266, doi.org/10.1016/j.cognition.2008.09.002.
- Burr, D. C., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*(6), 425–428.
- Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *The Journal of Neuroscience*, *32*(22), 7685–7700, doi.org/10.1523/JNEUROSCI.3325-11.2012.
- Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*(5), 401–406, doi.org/10.1111/j.1467-9280.2006.01719.x.
- Cantlon, J. F., Platt, M., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83–91, doi.org/10.1016/j.tics.2008.11.007.
- Cantrell, L., & Smith, L. B. (2013). Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition*, *128*(3), 331–352, doi.org/10.1016/j.cognition.2013.04.008.
- Castelli, F., Glaser, D., & Butterworth, B. (2006). Discrete and analogue quantity processing in the parietal lobe: A functional MRI study. *Proceedings of the National Academy of Sciences, USA*, *103*(12), 4693–4698, doi.org/10.1073/pnas.0600444103.
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404, doi.org/10.1016/S0042-6989(02)00596-5.
- Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*(1), 1–13.
- Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*(5), 408–411.
- Corbett, J. E., & Melcher, D. (2013). Characterizing ensemble statistics: Mean size is represented across multiple frames of reference. *Attention, Perception, & Psychophysics*, *76*(3), 746–758.
- Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences, USA*, *108*(49), 19552–19557, doi.org/10.1073/pnas.1113195108.
- Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neural model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407.
- Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 149.
- Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*, *18*(18), R855–R856.
- Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314, doi.org/10.1016/j.tics.2004.05.002.
- Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113*(1), 1–13.
- Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, *141*(4), 642.
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., . . . Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One*, *8*(6), e67374.
- Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. M. Wolfe & L. Robertson (Eds.), *From perception to consciousness: Searching with Anne Treisman* (pp. 339–349). New York: Oxford University Press.
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the “number sense”: The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–1465, doi.org/10.1037/a0012682.
- Halberda, J., & Odic, D. (2014). The precision and internal confidence of our approximate number thoughts. In D. C. Geary, D. Berch, & K. Koepke (Eds.), *Evolutionary origins and early development of number processing* (p. 305). Waltham, MA: Academic Press.
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334–337, doi.org/10.1038/383334a0.
- Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development*, *2*(2), 77–96.
- Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences, USA*, *106*(25), 10382–10385, doi.org/10.1073/pnas.0812142106.

- Joo, S. J., Shin, K., Chong, S. C., & Blake, R. (2009). On the nature of the stimulus information necessary for estimating mean size of visual arrays. *Journal of Vision*, 9(9):7, 1–12, doi:10.1167/9.9.7. [PubMed] [Article]
- Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *Proceedings of the National Academy of Sciences, USA*, 103(9), 3486.
- Kilian, A., Yaman, S., von Fersen, L., & Güntürkün, O. (2003). A bottlenose dolphin discriminates visual stimuli differing in numerosity. *Learning & Behavior*, 31(2), 133–142, doi.org/10.3758/BF03195976.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35(13), 1897–1916.
- Libertus, M. E., & Brannon, E. M. (2009). Behavioral and neural basis of number sense in infancy. *Current Directions in Psychological Science*, 18(6), 346.
- Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica*, 141(3), 373–379.
- 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(46), 18737–18742, doi.org/10.1073/pnas.1207212109.
- McPeck, R. M., Maljkovic, V., & Nakayama, K. (1999). Saccades require focal attention and are facilitated by a short-term memory system. *Vision Research*, 39(8), 1555–1566.
- Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(3), 320–334.
- Morgan, M. J., Raphael, S., Tibber, M. S., & Dakin, S. C. (2014). A texture-processing model of the “visual sense of number.” *Proceedings of the Royal Society B: Biological Sciences*, 281(1790), 20141137.
- Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *Journal of Experimental Psychology: General*, 43(1), 255–265, doi.org/10.1037/a0030825.
- Odic, D., Le Corre, M., & Halberda, J. (2015). Children’s mappings between number words and the approximate number system. *Cognition*, (138), 102–121.
- Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology*, 49(6), 1103–1112.
- Odic, D., Pietroski, P., Hunter, T., Lidz, J., & Halberda, J. (2013). Young children’s understanding of “more” and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(2), 451–461.
- Park, J., DeWind, N. K., Woldorff, M. G., & Brannon, E. M. (2015). Rapid and direct encoding of numerosity in the visual stream. *Cerebral Cortex*, bhv017.
- Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, 44(3), 547–555, doi.org/10.1016/j.neuron.2004.10.014.
- Piffer, L., Agrillo, C., & Hyde, D. C. (2011). Small and large number discrimination in guppies. *Animal Cognition*, 15(2), 215–221, http://doi.org/10.1007/s10071-011-0447-9.
- Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, 41(6), 983–993.
- Rensink, R. A., O’Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8(5), 368–373, doi.org/10.1111/j.1467-9280.1997.tb00427.x.
- Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. *Attention and Performance XV*, 15, 231–265.
- Roitman, J. D., Brannon, E. M., & Platt, M. L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biology*, 5(8), e208.
- Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision*, 12(4):14, 1–17, doi:10.1167/12.4.14. [PubMed] [Article]
- Ross, J., & Burr, D. C. (2010). Vision senses number directly. *Journal of Vision*, 10(2):10, 1–8, doi:10.1167/10.2.10. [PubMed] [Article]
- Ross, J., & Burr, D. (2012). Number, texture and crowding. *Trends in Cognitive Sciences*, 16(4), 196–197, doi.org/10.1016/j.tics.2012.01.010.
- Sengupta, R., Surampudi, B. R., & Melcher, D. (2014). A visual sense of number emerges from the dynamics of a recurrent on-center off-surround

- neural network. *Brain Research*, 1582, 114–124, doi.org/10.1016/j.brainres.2014.03.014.
- Stoianov, I., & Zorzi, M. (2012). Emergence of a “visual number sense” in hierarchical generative models. *Nature Neuroscience*, 15(2), 194–196, doi.org/10.1038/nn.2996.
- Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, 11(5):13, 1–82, doi:10.1167/11.5.13. [PubMed] [Article]
- Szucs, D., Nobes, A., Devine, A., Gabriel, F. C., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Frontiers in Psychology*, 4(444). Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3715731/>
- Tibber, M. S., Greenwood, J. A., & Dakin, S. C. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *Journal of Vision*, 12(6):8, 1–19, doi:10.1167/12.6.8. [PubMed] [Article]
- Viswanathan, P., & Nieder, A. (2013). Neuronal correlates of a visual “sense of number” in primate parietal and prefrontal cortices. *Proceedings of the National Academy of Sciences, USA*, 110(27), 11187–11192.
- Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, 19(9), 1395–1407.
- Weber, B., Schwarz, U., Kneifel, S., Treyer, V., & Buck, A. (2000). Hierarchical visual processing is dependent on the oculomotor system. *Neuroreport*, 11(02), 241–248.