Forthcoming in *Mind & Language*

The Big-Small Problem in Infant Number Cognition*

Sam Clarke (sam.clarke@usc.edu)
Department of Philosophy, University of Southern California

Abstract: When subitizing, infants precisely discriminate collections containing ≤3 items, after which performance falls to chance. It remains unclear, however, why performance falls to chance given that infants approximately enumerate larger collections. This is the big-small problem. This paper clarifies the problem, notes that it's exacerbated by influential ways of thinking about numerical cognition and argues that existing "solutions" prove unsatisfactory. It then develops an improved solution, which turns on independently motivated claims about mental formats and infant working memory. This improved solution has ramifications for numerical architecture, the structure of perceptual representations, and the ways in which perceptual states refer.

1. Introduction

Infants can *approximately enumerate* large collections, albeit imprecisely and in accord with Weber's Law, such that discriminations are ratio-sensitive (Izard et al. 2009; Xu & Spelke 2000). Infants can also *subitize*, or precisely discriminate small collections, but only when these contain less than ~3 items, after which performance falls to chance (Feigenson et al. 2004). What's unclear is *why* subitizing would consistently fall to chance under these conditions, given that infants possess the abovementioned ability to approximately enumerate larger collections. I call this *the big-small problem*. This paper aims to clarify the problem and then resolve it.

2. The Big-Small Problem

The big-small problem arises because human infants possess myriad numerical abilities, each distinguished by idiosyncratic signature limitations. This essay focusses on their abilities to

* Really, this was supposed to be a paper about the hip-hop legend, Biggie Smalls, but the editors (quite reasonably) put their feet down. For particularly detailed comments on an earlier version of this manuscript, I am grateful to Elizabeth Spelke and Sarah Beck. For helpful discussion, I am also grateful to Liz Brannon, Laurenz Casser, Susan Carey, Sara Cordes, Gabe Dupre, Victor Gomes, Alon Hafri, Tyler Knowlton, Kevin Lande, Toby Sandoval, as well as audiences at The University of Bristol, UC Davis, The University of Delaware, New York University, Ohio State University, and The Annual Interdisciplinary Society Conference in Wyoming.

approximately enumerate large collections and *subitize* small collections. In this opening section, I clarify these abilities, such that readers can appreciate an apparent tension between them.[1]

2.1 Approximate Enumeration

'Approximate enumeration' concerns a capacity to perceptually discriminate (sometimes quite large) numerical quantities without counting. This ability is imperfect. Specifically, accuracy conforms to Weber's Law: when we approximately enumerate two collections, the ease and reliability with which we discriminate these is predicted by the ratio between them rather than their absolute difference in number. Ten dots are easier to discriminate from eight dots than twelve, even though eight and twelve differ from ten by the same absolute amount. What matters is the ratio between the quantities – the further from 1:1, the better.

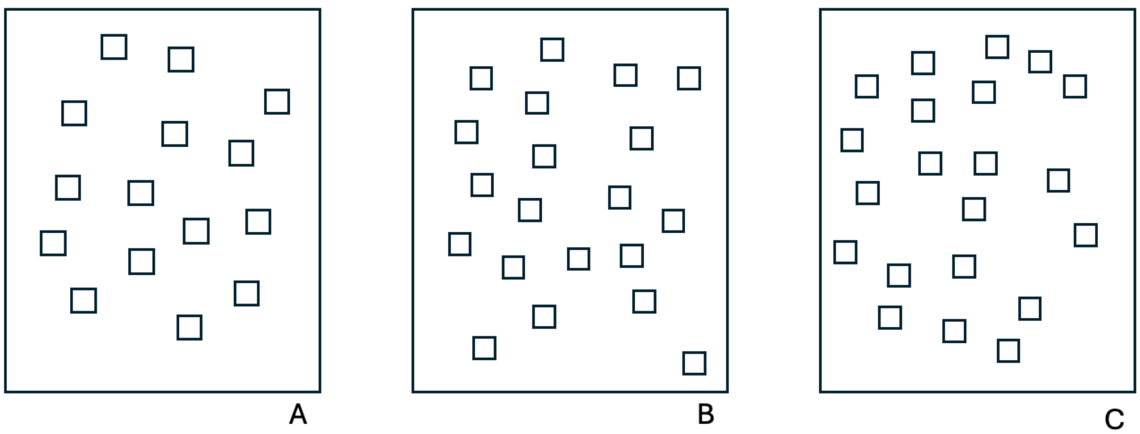

Figure 1. Based on a quick glance, can you tell whether panel A or panel B contains more squares? What about panels B and C?

Take Figure 1. Without explicitly counting, you will find it easier to recognize that panel B contains more squares than panel A than that panel C contains more squares than panel B. For instance, you

[1] Infants may possess additional capacities to precisely enumerate large collections in some contexts (Clarke 2025) and for approximate rational number representation (Denison & Xu 2014; Qu et al. 2024). These further capacities are orthogonal to the problem considered here, though related issues might arise when considering their interrelations.

Related questions may also arise when considering if/how the problem applies to non-human animals. If subitizing and approximate enumeration systems are evolutionarily ancient and preserved across species (Spelke 2003; 2023; Carey 2009), one might expect that the problem will arise (and its solution will apply) in much the same way as it does in human children. If, however, one takes the presence of an ANS, or ANS-like system, in such distant creatures as fish (Agrillo & Bisazza 2018) and insects (Giurfa 2019), or its diverging algorithms in nearby relatives like monkeys (Beran et al. 2024), to be evidence that approximate enumeration systems have emerged independently, many times over, through convergent evolution, we might expect the details of the problem and its solution(s) to differ. Given these potential complications, the present treatment focuses on the problem as it arises in human children but encourages consideration of these issues in future work.

might feel more confident that B>A than C>B or find yourself unsure as to whether C>B. The surface area, density, brightness, and convex hull of all three collections can be controlled for without affecting this basic result. The key difference is that the ratios between the numerical quantities varies, with the ratio of items in panels B to C closer to 1:1 than in panels A to B.

What's important for our purposes is that analogous capacities are found in young infants, even neonates (Izard et al. 2009; de Hevia et al. 2014). Here, researchers have taken care to show that performance involves a sensitivity to numerical quantities, rather than continuous properties of perceived collections (Clarke & Beck 2021). Moreover, diverse brain imaging methods indicate that these congenital abilities are underwritten by the same neural mechanisms employed by adults when they approximately enumerate (Cantlon et al., 2006, Hyde et al, 2010). In each case, performance is marked by a common signature limit – accuracy conforms to Weber's Law and is ratio-dependent, likening these infant abilities to those found in adult humans.

Take Xu and Spelke (2000). These researchers found that when six-month-olds habituated to collections of eight dots, they dishabituated to collections containing four dots or 16 dots, but not collections containing 12 dots. Likewise, six-month-olds who habituated to collections of 16 dots, dishabituated to 32 or eight-dot arrays, but not collections containing 24 or 12. Since non-numerical confounds, like surface area, density, and brightness were controlled for, it was concluded that the infants were enumerating the numbers of dots in each collection, but only approximately, such that discriminating two collections required them to differ in number by a ratio of at least 1:2.

Follow up studies have confirmed these results and shown that performance improves with age. For instance, Lipton and Spelke (2003) replicated the abovementioned findings in six-month-olds and found that performance was identical irrespective of whether infants discriminated seen dots or sequences of heard tones. They also found that nine-month-olds can perform harder 2:3 discriminations with comparable reliability. Both results replicate. Indeed, Libertus and Brannon (2010) replicated both results using a novel change detection paradigm and found that individual differences in task performance remained constant from one testing session to the next.

2.2 Subitizing

Where 'approximate enumeration' concerns a capacity to enumerate and discriminate *sometimes rather large collections* (e.g., collections of eight, 16 or 32 items) *imprecisely and in accord with*

Weber's Law, 'subitizing' or 'parallel individuation' concerns a capacity to *precisely discriminate small collections*, *up to a set-size of ~3 or 4*. [Those who deny that such abilities reflect explicit representations of number typically prefer the term "parallel individuation" to "subitizing" (Feigenson et al. 2004). Nothing I say turns on this, so I'll use the term "subitizing" for brevity.]

Like approximate enumeration, subitizing is observed in adults. W.S. Jevons (1871) found that humans are remarkably accurate at enumerating collections of 1-4 beans quickly tossed into a pan. More recently, researchers have confirmed these results, finding that humans spontaneously enumerate sets of ≤4 attended items quickly (40-100ms/item) and accurately, even when prevented from counting, while the enumeration of larger collections, outside this "subitizing" range, is slow (250-350ms/item) and error prone (Trick & Pylyshyn 1994). What's important here is that, yet again, analogous abilities are found in human infants – although, here, the set size limit is smaller (~3 rather than ~4) and, to reiterate, there is debate as to whether this involves infants explicitly representing numerical quantities (Feigenson et al. 2004; Margolis 2020).

Consider Feigenson et al. (2002). In this study, researchers found that 10-month-old infants would readily distinguish one item from two items, and two items from three items, but were at chance discriminating three items from four items, two items from four items, or three items from six items. A follow up study even found that they were at chance discriminating one item from four items (Feigenson & Carey 2005). So, while ten-month-olds reliably chose an opaque bucket into which two crackers had been sequentially placed over an opaque bucket containing one cracker, and an opaque bucket into which three crackers had been sequentially placed over an opaque bucket containing two crackers, they were at chance choosing between an opaque bucket containing six crackers and an opaque bucket containing three crackers or an opaque bucket containing four crackers and an opaque bucket containing just one. Bracketing the fact that infants reliably chose buckets with four crackers over empty buckets, it was as if the 10month-olds completely lost track of the quantities involved whenever one or both was larger than three.

Once again, these results replicate (see Wang & Kibbe 2024 for a recent meta-analysis). Similar results are found in manual search tasks (Feigenson & Carey 2003), tasks involving the discrimination of actions (Wynn 1996) or syllables (Bijelac-Babic et al. 1993), and in violation of expectation experiments (Wynn 1992). What's striking, is that performance in all these tasks falls apart when infants must discriminate non-empty collections containing >3 items, with the ratios among the collections failing to predict performance. Thus, 3:6 and 1:4 discriminations fail, while 2:3 discriminations succeed, despite 2:3's relative proximity to 1:1. This implies that performance in these tasks is not underwritten by infants' abilities to approximately enumerate. Instead, infants are seen to be drawing on a distinct psychological system, which facilitates *precise discriminations* but only among collections containing three items or less (Feigenson et al. 2004). In other words, performance in these tasks has been seen to implicate distinct subitizing mechanisms, operating according to their own psychophysical profile.

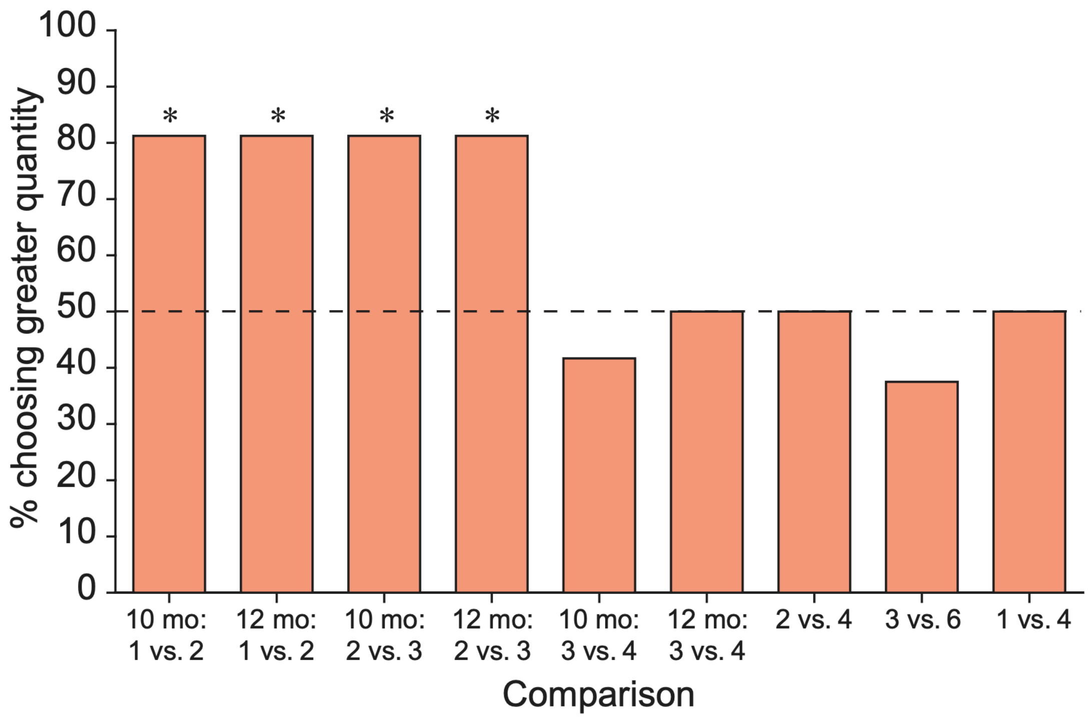

Figure 2. Feigenson et al. (2004) present the results of numerous subitizing tasks in which infants fall to chance when discriminating (non-empty) collections that contain >3 items.

2.3 The Problem

In short: *Approximate enumeration* concerns an ability to represent and discriminate (sometimes quite large) numerical quantities, albeit imprecisely and in accord with Weber's Law. Meanwhile *subitizing* concerns an ability to precisely discriminate small collections, but only when these contain less than ~3 items. But while both abilities are operational in young human infants, it's non-obvious how to reconcile their established signature limitations.

In studies evincing infant subitizing, infants succeed in precisely discriminating sets of one from sets of two and sets of two from sets of three. But recall: performance falls to chance when they're tasked with discriminating (non-empty) collections in which one or more contains >3 items. Thus, it's not that performance merely deteriorates when one or more collections contain >3 items. Rather, performance falls off a cliff (Figure 2). For instance, ten-month-olds tested by Feigenson et al. fell to chance when choosing between buckets containing two and four crackers (a difference of 1:2), or when choosing between buckets containing one and four crackers (a difference of 1:4). Indeed, this general tendency to catastrophically fail in subitizing tasks whenever (non-empty) set sizes exceed a threshold of three motivates the received view that subitizing is operational in human infants, and subject to its distinctive set-size limitations (Feigenson et al. 2004).

What's puzzling is that infants can approximately enumerate larger collections. For instance, six-month-olds readily discriminate collections of four from eight, collections of eight from 16, or other quantities that differ by at least 1:2 (Xu & Spelke 2000; Libertus & Brannon 2010). This is so, even when tested on a small number of trials – when more trials are run, harder ratios can be discriminated above chance (Halberda 2016). Indeed, such capacities are present at birth (Izard et al. 2009) and increase in acuity throughout development (Libertus & Brannon 2010). Accordingly, the 10month-olds tested by Feigenson et al. (2002) should have been *better* approximate enumerators than the six-month-olds tested by Xu and Spelke (2000). But given this capacity to approximately enumerate larger collections, outside the subitizing range, why would performance fall to chance in subitizing tasks whenever one or other collection contains >3items? Even if there is a three-item limit on what can be handled by infants' subitizing system, and this is explained by the system's basic architecture or algorithms (Feigenson & Carey 2005), why wouldn't infants discriminate small collections from large collections outside the subitizing range, imprecisely yet significantly above chance, by approximately enumerating them, at least when the collections differ by a suitably large ratio – e.g., 2:4 or 1:4? This is *the big-small problem*.

2.4 Clarifying the Problem

To clarify the problem, it's instructive to consider a simple response to it, which may sound tempting, but proves unsatisfactory. This simple response starts from the observation that studies of approximate enumeration often involve participants enumerating collections of items that are presented *concurrently* or *all at once*. For instance, in work by Xu and colleagues, six-month-olds were habituated to collections of eight or 16 dots, where all the dots were presented simultaneously on a screen. This leaves open that the mechanisms of approximate enumeration might be unable to enumerate items presented sequentially, one-by-one. Since the 10-month-olds in Feigenson et al.'s studies observed crackers being placed one-by-one into opaque buckets, you might then think that the reason why non-subitizable values failed to be approximately enumerated (and, thus, discriminated) is because approximate enumeration requires the simultaneous presentation of *all the elements* being enumerated (see Spelke 2003: 298 for a loosely related suggestion).

This simple suggestion might even seem to be well motivated. Wynn (1995) found that five-month-olds subitize small collections that are sequentially placed behind a screen. For when infants saw two objects get sequentially placed behind a screen, they looked longer when the screen was removed to reveal three objects rather than two, despite marked difficulties five-month-olds have performing 2:3 discriminations by approximately enumerating. However, in a similar study with older children, Chiang and Wynn (2000) found that when eight-month-olds saw five objects get sequentially placed behind a screen, in analogous ways, they failed to approximately enumerate these and were no more surprised when the screen was removed to reveal five objects or none. Taken together, these findings may indicate that while infants approximately enumerate collections containing >3 items, this requires that items be presented simultaneously rather than sequentially.

Alas, matters are not so simple. Infants *can* approximately enumerate collections when items are presented one-by-one, and sequentially, at least under some conditions. In fact, we've already seen this. When Lipton and Spelke (2003) replicated Xu and Spelke's (2000) results with six-month-olds, performance was identical when infants enumerated *sequences* of heard tones as opposed to static collections of seen dots. Likewise, Wood and Spelke (2005) found that both six- and nine-month-olds could approximately enumerate sequences of rabbit jumps, again with comparable levels of acuity to when approximately enumerating static collections. Infants also show a three-item limit in subitizing tasks when items are presented concurrently (Starkey & Cooper Jr. 1980).

This is not to deny that there is *something* about the way collections are presented which determines whether infants approximately enumerate, and thus discriminate, these when set size is >3. Spelke (2022: 156) conjectures that this often has to do with the amount of time that items are presented for. Specifically, she proposes that while infants approximately enumerate collections quickly and in parallel, irrespective of whether items are presented sequentially or concurrently, subitizing requires that arrays be presented slowly enough that each item is selectively attended. As evidence for this, Spelke cites studies in which collections containing ≤3 items are presented briefly, precluding such focused attention, and are approximately enumerated and, thus, discriminated in accord with Weber's Law (Starr et al. 2013; Libertus et al. 2014).

The point to note is that even if this is so, it does not resolve the big-small problem. Suppose Spelke's hypothesis about presentation times is correct: subitizing requires that individual items be selectively attended while approximate enumeration does not; and because selectively attending to individual items takes time, infants have no choice but to approximately enumerate small collections when presentation times are brief – subitizing is simply impossible under these conditions. Fine. The trouble is: While this might explain why infants would sometimes resort to approximately enumerating small collections, it does not explain why infants would not approximately enumerate collections containing >3 items they would otherwise subitize (e.g., in Feigenson et al.'s 2002). Think about it this way: If infants can flip from approximately enumerating items to subitizing them when set-size is small and items are presented for a sufficient duration that individual items can be selectively attended, why can't they flip back to approximately enumerating larger collections when a set-size of 3 is exceeded and subitizing fails? Indeed, if Spelke's hypothesis about presentation times is correct, this arguably exacerbates matters. For if approximate enumeration is so easy that it operates quickly and without even requiring focused attention to enumerated individuals, it only looks more puzzling that this fast and efficient ability cannot save the day in subitizing tasks when a set-size of three is exceeded, especially when all the items are presented simultaneously (Starkey & Cooper Jr. 1980).

3. Mutual Inhibition

The preceding remarks introduce the big-small problem. Boldly stated, we want to know why infants fall to chance in subitizing tasks whenever one or more (non-empty) target collection contains >3 items, even when these collections differ by a ratio that should be easily discriminable via their well-known competence to approximately enumerate. What gives?

Perhaps the most prominent answer appeals to some form of inhibition. Exemplifying this view, Spelke (2022) proposes that approximate enumeration and subitizing are subserved by distinct modules, in roughly Fodor's (1983) sense of the term: there is, thus, a modular approximate number system (ANS) and a distinct subitizing module, operating according to its own proprietary data structures and algorithms. But, drawing on work by Hyde and Wood, she adds that these systems are unlike sensory modules in that they are "mutually inhibitory" (2022: 170). So, while the ANS and subitizing systems are distinct psychological systems, Spelke proposes that activation of the ANS or subitizing system inhibits that of the other. Such inhibition is then seen to defuse the big-small problem, providing an "account for all the puzzling findings" (169) under consideration.

Spelke says little to clarify what 'inhibition' amounts to in this context. Nor do other proponents of this general suggestion (Hyde 2011). But, for our purposes, this won't do. If we are to resolve the problem in this way, 'inhibition' must amount to more than (e.g.) an illicit re-labeling of the fact that the ANS fails to facilitate discriminations in subitizing contexts. With this in view, the present section will pose problems for two natural ways of fleshing out Spelke's inhibitory hypothesis, thereby motivating consideration of my recommended alternative.

3.1 Strong Inhibition

While Spelke says little to clarify what 'inhibition' amounts to in this context, she sometimes seems tempted by a view which I call *Strong Inhibition*. On this view, activation of the subitizing system straightforwardly switches off the ANS. Therefore, when infants are tasked with choosing between a four-cracker-collection and a one-cracker-collection they will be at chance if their subitizing system has already been activated, since their subitizing system cannot process the four-cracker-collection (this exceeds its three-item limit) and activation of the subitizing system turns off the ANS – a system which could have otherwise handled the 1:4 discrimination. Since similar points apply whenever one or both collections exceed a set-size limit of three, Strong Inhibition offers to defuse the big-small problem.

Indicative that this is what Spelke has in mind, consider that her main line of evidence in support of the inhibitory hypothesis comes from Hyde and Wood (2011). Building on prior studies showing that collections can be approximately enumerated, even when positioned such that their individual constituents cannot be selectively attended (Intriligator & Cavanaugh 2001), and prior studies finding common set-size limitations in subitizing and object-based-attention tasks (Trick & Pylyshyn 1994), Hyde and Wood took EEG measures while adult subjects viewed small collections, containing 1-3 items, which either could or couldn't be allocated selective attention. They found that when objects were sufficiently spaced and foveated, such that individual items could be allocated focused attention, EEG responses were observed that are standardly associated with subitizing. This was taken to suggest that such attentional deployment automatically elicited subitizing. Meanwhile, when individual objects could not be selectively attended, responses were comparable to those evoked in studies probing the neural underpinnings of the ANS. But crucially, when participants selectively attended to the individual objects, and evoked EEG responses associated with subitizing, Spelke (2022: 168) emphasizes that this was accompanied by "no detectable neural response to changes in number" of a sort one would expect if participants were engaging their ANS. It was as if attending to individual items, in a manner that automatically evoked subitizing, switched off the ANS, just as Strong Inhibition recommends. Indeed, the lead author of the cited study invites this interpretation, concluding that when subitizing is engaged "approximate number representations are not formed" (Hyde 2011: 4).

Strong Inhibition might, therefore, seem to have a lot going for it. It offers to defuse the big-small problem, and it does so in a manner that is supported by sophisticated neuroscientific results. Despite these virtues, I think Strong Inhibition untenable.

If selectively attending to individuals, and subitizing them, literally switched off one's ANS, as Strong Inhibition recommends, it should be impossible to approximately enumerate the total number of items populating an array when we selectively attend to a small number of individuals within this larger array and subitize them. But this does not seem to be so.

Let me pump your intuitions: Consider Pylyshyn's Multiple Object Tracking (MOT) paradigm. In standard MOT experiments, participants are presented with non-subitizable collections of items (e.g., 10 black dots). At the start of each trial, a subset of these is flagged as 'targets' to be tracked throughout the experiment – for instance, target dots might flash on the screen. Having stopped flashing, all the items in the array begin moving in unpredictable ways for, say, 10secs. At this point, the items freeze, and participants are tasked with reidentifying the targets highlighted at the start of the trial.

A much-celebrated result is that adult humans typically succeed in tracking up to 3 or 4 targets in such tasks after which performance falls apart (Pylyshyn 2007). This has been widely noted to mirror the set-size limitations on subitizing (Scholl & Leslie 1999; Feigenson 2011; Spelke 2022). And sure enough: participants in these tasks have been found to spontaneously subitize and enumerate the items being tracked (Trick & Pylyshyn 1994). But if you try a MOT study for yourself, notice that when you track and enumerate 3-4 target objects, through focused object-based attention towards these, you are not left oblivious to the approximate number of items populating the entire array. In fact, you might feel that you can't help but notice this. For instance, you might be unsure whether the collection contains precisely ten or precisely 12 dots but be confident that it doesn't comprise 20. This would presumably be so, even if your attention were focused on the target items *before* distractor items were even inserted into the scene – something which would preclude an ANS representation being formed for the entire collection (and subsequently memorized) *prior* to engagement of the subitizing system.

This is not mere conjecture. Available evidence indicates that when selective attention is allocated to individual items within a collection this *improves* ANS acuity. For instance, Cheyette and Piantadosi (2019) used an eye tracker to show that the more dots that participants could selectively fixate upon within a collection, the more accurately they could approximately enumerate the collection. So, when participants selectively attend to individuals within a collection, in ways that are seen to elicit subitizing (Trick & Pylyshyn 1994; Spelke 2022), this does not suppress their ability to approximately enumerate that very collection, as Strong Inhibition predicts – on the contrary, it *improves* this, iteratively refining the acuity of ANS representations.

With these points in view, it is worth revisiting Hyde and Wood's study, noting that in hindsight it did not directly test for the inhibition of the ANS by the subitizing system. For a start, it employed a debated form of "reverse inference" (Poldrack 2006). For even if activation of the subitizing system and its underlying *neural machinery* was shown to inhibit *neural machinery* associated with the ANS (a point which may, itself, prove difficult to assess given limitations of EEG studies [Grech et al. 2008; Srinivassan 1999; but see Gennari et al. 2023 for recent methods which address some of these concerns in the context of infant numeracy), Hyde and Wood's study did not test whether participants retained an ability to approximately enumerate collections when subitizing; e.g., by asking them to concurrently estimate or discriminate a collection of dots whilst subitizing was engaged. Of course, in the absence of contravening evidence, Hyde and Wood's results might motivate the thought that they would not (see: Machery 2014). But available behavioral results seem to trump this suggestion. Thus, I submit that Strong Inhibition should be rejected. It is undermotivated by the studies that are seen to support it, and it is undermined by other available evidence (at least in adult subjects, akin to those tested by Hyde and Wood).

3.2 Weak Inhibition

Strong Inhibition seems to be too strong. Nevertheless, one might embrace a modest version of the inhibitory hypothesis, which I call *Weak Inhibition.* Given Weak Inhibition, activation of the subitizing system does not straightforwardly switch off the ANS, but it reduces ANS acuity, causing performance to "suffer" (Spelke 2022: 169). In the tasks under consideration, this might prevent infants' ANSs from facilitating discriminations we would otherwise expect them to.

One problem with this proposal is that even Weak Inhibition is undermined by the above studies. For when adults deploy focused attention towards individuals within a collection, we have now seen that this *improves* ANS acuity, rather than reducing it (Cheyette & Piantadosi 2019). Insofar as such focused attention automatically elicits subitizing (as Spelke [2022: 169] and Hyde & Wood [2011] maintain), and the system works in broadly homologous ways across development (Carey, 2009; Clarke et al. 2025), this calls even Weak Inhibition into question.

Bracketing these concerns, Weak Inhibition might find support in developmental studies. Recall Spelke's suggestion that while the ANS operates rapidly (enumerating collections before focused attention can be paid to the individuals these comprise), subitizing often requires that items be observed for longer periods of time (such that focused attention can be paid to individual items within a collection). As evidence for this, Spelke notes that when small collections containing ≤3 items are only presented briefly, they are often approximately enumerated, imprecisely and in accord with Weber's Law, rather than being precisely subitized (Brannon 2002; Wood & Spelke 2005). What she neglects to mention, is that when infants compare subitizable and non-subitizable quantities in these studies, the ratios between these collections must be larger for reliable discriminations to obtain. For instance, Cordes and Brannon (2009a; 2009b) found that when seven-month-olds approximately enumerated and proceeded to distinguish subitizable quantities from non-subitizable quantities, these quantities needed to differ by a ratio of 1:4, under conditions where the discrimination of two non-subitizable quantities otherwise required a mere 1:2 difference. *Prima facie*, ANS acuity was (roughly) halved when discriminating collections that crossed the subitizing threshold, perhaps due to engagement of the subitizing system (but note: Cordes & Brannon 2009a employed a habituation paradigm, effectively ensuring that collections were presented for prolonged periods, which should activate the subitizing system on Spelke's hypothesis about presentation times).[2] So, while activation of the subitizing system may not switch the ANS off (*pace* Strong Inhibition), this suggests that ANS acuity is lowered when subitizable collections are involved and the subitizing system is engaged (*ibid.*).

What's crucial for us is that inhibition of this weakened sort no longer resolves the big-small problem. Since infants continue to consistently perform 1:4 discriminations across the subitizing threshold using their ANS, despite prolonged presentation times, it remains unclear why infants consistently failed to discriminate one cracker from four crackers in (e.g.) Feigenson and Carey (2005) using their ANS. Indeed, this point is exacerbated when we consider that the infants in Feigenson's studies were older (ten-months-old) and would, thus, be expected to discriminate harder ratios than the seven-month-olds tested by Cordes and Brannon (see: Libertus & Brannon 2010). While there might be other ways to flesh out Weak Inhibition, these concerns motivate a fresh approach to resolving the big-small problem.

4 An Independently Motivated Solution

I have raised concerns with two formulations of the inhibitory hypothesis. I'll now suggest that a neglected solution to the big-small problem presents itself when we avail ourselves of three independently motivated observations about the ANS and subitizing system that more-or-less all parties in this debate already accept. That:

a) these systems' outputs differ in format,
b) these outputs compete for space in visual working memory
c) there are cues *of some sort* which determine whether a small collection is discriminated by the subitizing system or the ANS.

Let us consider these claims in turn.

4.1 Format

In discussions of number cognition, it is widely accepted that approximate number representations are couched in an *analog* format. In saying this, we needn't assume that these representations are continuous (Beck 2015; c.f. Gallistel & Gelman 2000), nor that they conform to Kosslyn's (1980) "picture principle" (Clarke 2022; c.f. Carey 2009). What's crucial is that number is represented by

---

[2] This might be a slight overstatement, since Cordes and Brannon did not test 1:3 discrimination across the subitizing threshold (e.g., 2:6). I'll put this to one side since acknowledging this only worsens the problems for Weak Inhibition.

a magnitude in the head (Peacocke 2019), which functions to mirror the quantities being represented, varying as a monotonic function of these (Maley 2011; Beck 2015). To conceptualize this, approximate number representations are thought to be like the analog representations in mercury thermometers which represent temperatures by having mercury levels vary as a monotonic function of these (i.e. by having mercury levels serve as *analogs* of their content).

Beyond the fact that this analog format is arguably evinced by neuroscientific results in closely related, non-human primates (Roitman et al. 2007; Nieder 2016), a key motivation for positing analog representations in this context, has been the observation that *if* the ANS's representations are couched in an analog format, this could explain the ANS's conformity to Weber's Law given how noise naturally accumulates in analog systems (Meck & Church 1983).

To appreciate the basic idea, imagine keeping count of goals scored at a football match by pouring one cup of water into a bucket A when team A scores a goal and a separate cup of water into a bucket B when team B does. Under these conditions, you might expect the bucket with the most water at the end of the game to correspond to that of the winning team. However, if each cupful varies in volume, noise of this sort could lead to systematic "errors" that conform to Weber's Law. For instance, if each cupful poured into bucket A contains 200ml of water and each cupful poured into bucket B contains 300ml of water, then the bucket with the most water will reliably correspond to that of the winning team just in case the winners win by 2 goals to 1, but not by 3 goals to 4, irrespective of the absolute numbers of goals involved. More generally: accuracy will now be predicted by the ratio between the numbers of goals scored. Since noise is inevitable in real-world analog systems, like those implemented in the brain, a prevailing orthodoxy has been that Weber's Law results from the analog format of the representations underwriting ANS performance plus the patterns of noise that arise in the brain's construction of its sensory representations (see Beck & Clarke *forthcoming* for more detail).

What's matters for us is that this should compel us to hold that subitizing encodes numerical information in a different format entirely. For if one holds that the format of ANS representations implies the ANS's conformity to Weber's Law, then the fact that subitizing does not conform to Weber's Law implies that its representations must somehow differ in how they make numerical information "explicit" and "accessible" (Marr 1980: 20-22). For instance, if the subitizing system represents small numbers precisely (Margolis 2020), one might conjecture that these are couched in a non-analog digital format – e.g., that these are digital symbols, ONE TWO and THREE, in the language of thought (Quilty-Dunn et al. 2023). Alternatively, one might deny that subitizing involves the explicit representation of numerical content and instead hold that the discriminations performed in subitizing tasks result from numerical information that is merely implicit in the number of individuals explicitly represented (Feigenson et al. 2004). Either way: Subitizing representations cannot encode numerical information in the same format as approximate number representations if approximate number representations have a format which implies the ANS's conformity to Weber's Law. This is for the simple, and uncontroversial, reason that subitizing does not conform to Weber's Law.[3] Thus, I draw **Interim Conclusion 1 – ANS and subitizing representations encode numerical information in different formats.**

4.2 No Comparison Without Translation

While Interim Conclusion 1 is relatively uncontroversial, at least among those engaged with the big-small problem (e.g., Feigenson et al. 2004; Spelke 2022; Carey 2009; Brannon 2002; see also: Gallistel 2011), it suggests that the numerical information that the ANS and subitizing systems encode cannot be directly compared without some intervening process of translation. Just as one will not be able to identify which of two symbols represents a larger number if one number is represented using a system of tally marks and the other using Arabic numerals (unless, of course, one possesses the knowledge or ability to translate these into a common code) the numerical information that ANS and subitizing representations carry will be incomparable without some capacity to translate this information into a common format. This is particularly clear if – as is often assumed – computations over these representations must ultimately be sensitive to their syntactic, or non-semantic properties (Fodor 1979). Thus, I proceed to draw **Interim Conclusion 2 – If [Interim Conclusion 1] is accepted, then the contents of ANS and subitizing representations cannot be contrasted without an intervening process of translation.**

[3] This point applies, even if one rejects the conjecture that ANS representations are analog and endorses its chief rival in the extant literature: namely, Gallistel's (2011; 2017) "autoscaling" hypothesis. According to this rival hypothesis, magnitude representations have a "bipartite" format, "like the representation of quantity in scientific notation" where "one part specifies where the magnitude falls within some range; the other specifies the range (scale)." Like the analog hypothesis, this is argued to imply and explain the emergence of Weber's Law in a system's discriminations (but see Beck & Clarke *forthcoming* for critique). But, for this reason, my basic point still holds: ANS representations cannot encode numerical content/information in the same format as subitizing representations, since subitizing representations are not susceptible to Weber's Law.

4.3 Resolving The Problem

Once again, Interim Conclusion 2 is largely uncontroversial (see Butterfill & Sinigaglia [2014] and the literature it's spawned), but it has implications for the big-small problem. For if the contents of the ANS and subitizing system differ in format (Interim Conclusion 1), we can now see that an infant in Feigenson et al.'s studies will not be able to directly compare the quantities associated with their representation of a four-cracker collection and their representation of a one cracker collection if the former has been represented by their ANS (in one format) and the latter representation has been encoded by their subitizing system (in another) – at least not without an extraneous process of translation (Interim Conclusion 2) which we have little reason to think infants possess.[4] Indeed, similar points apply no matter what these representations represent (compare Feigenson et al. 2004; Margolis 2020; Clarke & Beck 2021; Samuels & Snyder 2024), and – hence – no matter the ratio between the collections being referenced.

Simply saying this does not resolve the big-small problem, however. The preceding points allow us to see why an approximate number representation of FOUR (or FOURISH) and a subitizing representation of a one item collection will be incomparable for infants in Feigenson et al.'s studies. Nevertheless, an appeal to the diverging formats of approximate enumeration and subitizing does not explain why infants would systematically fail to discriminate four crackers from one cracker in Feigenson and colleagues' experiments. This is because it does not explain why infants in such studies wouldn't construct approximate number representations of both quantities – ONE (or ONEISH) and FOUR (or FOURISH), respectively – in a common ANS format and proceed to compare *these*. After all, we've seen that small collections *can* be approximately enumerated (e.g., Brannon 2002; Cordes & Brannon 2009a; 2009b).

Fortunately, a solution to the problem presents itself when the preceding remarks are considered in tandem with two final observations:

Firstly, utilizing ANS and/or subitizing representations requires working memory. No one doubts this – i.e., that there are limits on the amount of visual information that can be stored and accessed by an infant at a given time for use in reasoning or action guidance. What's notable is that, as a

---

[4] This contrasts with adults, who plausibly overcome The Big-small problemby mapping the outputs of subitizing (Trick & Pylyshyn 1994) and approximate enumeration (Sullivan & Barner 2013) onto lexical number concepts, couched in a common format, apt for immediate comparison (see Spelke 2003 for a related conjecture).

matter of empirical fact, current research indicates that infants are limited to holding no more than ~3 objects/collections in working memory in a range of tasks that are relevant to our concerns in this essay.[5] For instance, infants typically only succeed in tracking and retaining information about ≤3 Spelke objects at any given time (Feigenson & Carey 2003). Likewise, in approximate enumeration and subitizing tasks, infants can only compare information about ≤3 collections simultaneously, after which performance sharply declines. For instance, Zosh et al. (2011) found that infants detect changes in number that pertain to either one of two observed subsets or the superset they comprise, but they lose track of this when more than two subsets plus a superset are tracked simultaneously. Moher and Feigenson (2011) extended these results, showing that this three-collection-limit remains constant, irrespective of whether the subsets of a collection are demarcated by color or shape. Meanwhile, Halberda et al. (2006) found related capacity limits in adults, with related results observed in the subitizing literature. For instance, Feigenson and Halberda (2004) showed that infants can exceed the 3item set-size limit in subitizing tasks when collections are easily chunked into less than 3 subitizable sets (e.g., such that a collection of four can be represented as two collections of two).

In each of these cases, the representations involved seem to possess an object-attribute structure (Clarke 2023). Each representation is complex, picking out an individual (be it an isolated Spelke object, as in classic work on object files [Green & Quilty-Dunn 2021], or a collection apt to be subitized or approximately enumerated [Feigenson 2011]) such that information in various formats can then be attributed to these individuals (be it information about the kind of object being referenced, the individuals it comprises, or the approximate number of items that it contains – *ibid.*). Beyond the fact that this complex object-attribute structure allows that distinct types of information be bound and updated with respect to a single individual (e.g., average dot size *and/or* approximate number), indicating that the collection is not defined for the visual system by any of these specific attributes (Clarke 2023; compare Pylyshyn 2003), this conjecture is motivated by

[5] There are different views on the architecture of working memory. Some regard it as a horizontal faculty, such that (e.g.) *verbally encoded information* and *visually encoded information* compete for space within a single finite resource (Atkinson & Shiffrin 1968). Others argue for the functional independence of (e.g.) *visual* and *verbal* working memory, holding that each type of working memory stores content in an independent memory store (Shah & Miyake 1996). Still others endorse the functional independence of the memory stores associated with (e.g.) *visual* and *verbal* working memory but hold that memory storage within either store is constrained by domain general resources, common to visual and verbal cognition (Baddeley 2000; Kane et al. 2004). I take no stand on these matters. What's important for my argument is simply that subitized visual information competes with approximately enumerated visual content for space within a single finite memory store. All the abovementioned views predict this.

the three-item limit described above. For this pertains (most immediately) to the number of objects/collections that can be visually referenced at a given moment. Thus, it is as if there are just three slots available in infants' visual working memory, and each slot is clogged up whenever the infant thinks about an item – be it a Spelke object, a subitizable collection, or an approximately enumerated collection (Feigenson et al. 2011) – irrespective of what information (if any) is then attributed to this. Thus, I draw **Interim Conclusion 3 – Infants can hold information about no more than three collections in working memory at once.**

This, alone, still does not resolve the big-small problem. Consider the case in which an infant who discriminates two crackers from three crackers, falls to chance discriminating a one-cracker collection from a four-cracker collection. As noted, this is puzzling. For even if the subitizing system is unable to attribute the relevant numerical information to the four item collection (because the infant subitizing system has a three-item limit on the collections it can quantify or process), it remains unclear why infants could not use their ANS to perform the comparison, producing a representation of approximately one cracker for one collection which is then easily discriminated from an approximate number representation of four(ish) crackers in the other – after all, 1:4 is miles from 1:1.

The point to note is that, given Interim Conclusion 3, there simply will not be space for infants to simultaneously encode (i) an approximate number representation of the one-cracker collection, (ii) an approximate number representation of the four-cracker collection, (iii) a subitizing representation that refers to the one-cracker collection, and (iv) a subitizing representation that refers to the four-cracker collection. Only three of these four representations will be able to fit into the available slots that infant visual working memory provides. But which three?

At this point, it is worth reminding ourselves that there simply must be cues which reliably determine whether small collections are discriminated using the ANS or subitizing system. Without these it is hard to see why infants would consistently subitize in Feigenson et al.'s tasks, yet consistently approximately enumerate otherwise similar collections in studies conducted by the likes of Cordes and Brannon.

Admittedly, we do not have a clear understanding of what these cues are. Nevertheless, plausible proposals have been advanced. For instance, Spelke thinks this often has to do with presentation times: when collections/items are presented briefly, infants will tend to approximately enumerate these in accord with Weber's Law (and thus discriminate one item from four, just as reliably as they discriminate two items from eight). Meanwhile, she thinks longer presentation times dispose infants to subitize the collections (hence why they fail to discriminate collections that exceed the subitizing threshold under these conditions, as when they fail to discriminate one item from four [Feigenson & Carey 2005]).

Admittedly, Spelke's appeal to presentation times can't be the whole story. For a start, it doesn't explain why infants sometimes approximately enumerate small quantities in habituation studies (such as those employed by Cordes & Brannon 2009a) which typically involve longer presentation times than the subitizing studies run by Feigenson and colleagues (S. Cordes pers. comm.). It also doesn't explain why Cordes and Brannon consistently found that ratios need to be significantly larger (e.g., 1:4 rather than 1:2) for seven-month-olds to discriminate collection sizes that cross the subitizing threshold (but see Section 3.2) – a point which leads Cordes and Brannon (2009a) to suppose that ratio size may be a further cue that disposes approximate enumeration over subitizing in the tasks under consideration. In any case, it's important to stress that whatever cues end up determining performance in these studies, they need not be construed as cues which determine which system is activated in each context and which is switched off (*pace* Strong Inhibition). Instead, these cues can be seen to determine which systems' representations are prioritized in the three slots that infant working memory provides. Thus, when items are presented slowly, or in ways which otherwise dispose collections to be subitized, we might suppose that subitizing representations are just preferentially encoded into working memory over approximate number representations of the same collections. This does not require us to posit any further sense in which the ANS and subitizing systems suppress or inhibit one another's operations. This is a welcome result, I think, since alternatives run afoul of the concerns glossed in Section 3.

To see how this offers to resolve the big-small problem, consider Feigenson et al.'s cracker experiments one last time. In these studies, infants' subitizing systems will have tried to represent the collections and, due to the ways in which the items were presented, these will have been preferentially encoded into two of three available slots in infants' working memory. Thus, when infants observed two crackers being placed into one bucket and three crackers being placed into the other, a subitizing representation of the two-cracker collection and a subitizing representation of the three-cracker collection will have been stored in working memory, occupying two of the three available slots therein. Since both collections contain a subitizable quantity of items, this would have enabled the infants to then discriminate the collections and reliably choose the three-cracker collection over the two-cracker collection.

By contrast, consider a situation in which infants observed one cracker being placed into one bucket and four crackers into the other. Here, they would fail to discriminate these quantities. Provided that these collections were presented in analogous ways, the subitizing system would still try to represent the collections and these representations would still be preferentially encoded into working memory. These subitizing representations would, thus, continue to occupy two of three available slots therein. However, in this case, the subitizing system would be unable to attribute the information that is required to encode the quantity associated with the four-cracker collection (four-crackers exceeds its three-item limit). Thus, these representations would not enable a 1:4 discrimination. But while the ANS *might* have stepped in to facilitate this discrimination (given the large ratio size), the prioritization of subitized content into two of the three available working memory slots would only leave one additional slot available. So, even if this remaining slot allowed for an approximate number representation of the otherwise indiscriminable four-cracker collection to be encoded and utilized by the infants, this single representation would not suffice to facilitate a comparison among the collections if its numerical content were couched in a distinct format; at least not without some intervening means of translation (Interim Conclusions 1 & 2).

Since similar points apply to all the problem cases under consideration – including cases in which infants must discriminate between a three-cracker collection and a four cracker collection, a two-cracker collection and a four cracker collection, or a three-cracker collection and a six cracker collection – I propose that the big-small problem is resolved when we recognize that (i) the ANS and subitizing systems encode numerical information in different formats, (ii) their representational outputs compete for space in working memory, and (iii) there are cues determining whether discriminations are facilitated by the ANS or subitizing systems in these tasks, where (iii) is interpreted as a claim about what information is *prioritized* in working memory.

4.4 An Objection

One worry with my proposal might be that it requires us to hold that when a >3 item collection is presented in ways that cue subitizing representations to be preferentially encoded in working memory, a representation of the >3 item collection will somehow remain stored and prioritized in working memory (clogging up one of the available slots) over and above usable approximate number representations of the same collection. This could sound bizarre if one assumes that the subitizing system stops representing the collection *entirely* whenever that collection exceeds the subitizing threshold. For, on this view, there will be no such thing as an infant's subitizing representation which represents a four-item collection. When the subitizing threshold of three is crossed, the representation simply ceases to exist.

This cannot be the right way to think about subitizing, however. If a subitizing representation that refers to a four-cracker collection was not held in working memory *at all*, infants would reliably select buckets containing one cracker over buckets containing four. Why? Because the subitizing system would represent the one cracker bucket as containing one cracker while failing to represent the four-cracker collection *at all*. But, as we've seen, this prediction is not borne out (Feigenson & Carey 2005). Infants are at chance choosing between subitized collections of one and four.

You might wonder why this would this be. What would cause infants to be at chance choosing between collections of one and four in these studies? The answer, I suggest, stems from my prior suggestion that both subitizing and ANS representations have an object-attribute structure. Like object-files, which pick out perceptual objects and attribute information to these, these systems pick out *collections* of items before allowing syntactically independent symbols or representations (carrying numerical information) to be attributed to these (either explicitly [Margolis 2020; Clarke & Beck 2021] or implicitly in the number of individuals these collections are represented as containing [Feigenson et al. 2004]). But since the structure of these representations is complex, and the three available slots in infant visual working memory is clogged up by the elements of these representations which pick out objects or (in the cases at hand) collections (rather than the information that is then attributed to these) we should expect a subitizing representation to continue to consume a slot in visual working memory when the subitizing threshold is exceeded. Why? Because, under these conditions, there will still be a collection being picked out; it's just that the subitizing system will not be able to attribute the relevant numerical information to this object. In this way, the element that is stored in working memory will effectively say "there is a collection there" but no quantitative information will be attributed to it. So, when this gets compared to a separate subitizing representation which says "there is a collection there and it contains one item" these representations will underdetermine which collection contains more. By contrast, when a desirable four-item collection is picked out by an object-specifying element which says "there is a collection there" this collection will be chosen preferentially over an empty container, which infants have no reason to have ever treated as collection-involving to begin with. Consistent with this prediction, Feigenson and Carey (2005) found that while 10month-olds fail to perform 1:4 discriminations in subitizing tasks, they reliably perform 0:4 discriminations under comparable conditions, choosing a four-cracker collection over a bucket that is left empty.

5 Future Directions

This paper has clarified the big-small problem and recommended a solution to it. My solution is simple in that it turns on independently motivated claims that are already accepted by most parties in these debates. The basic idea is that, given the limited number of slots that are available in infants' visual working memory there is not space to encode two approximate number representations (pertaining to collections in and outside the subitizing range) if visual cues lead the visual system to preferentially encode two subitizing representations in working memory – this would require four slots, where infant working memory provides three. And while the storage of two subitizing representations might leave one slot free for an ANS representation to squeeze in, this won't facilitate a content-respecting comparison with the contents of subitizing representations stored in working memory if these are couched in different formats. This is because, content-respecting comparisons across formats require a (*prima facie* lacking) mechanism of translation.

This conjecture is at the mercy of empirical fortune. Since my hypothesis predicts that ANS and subitizing representations compete for space in working memory, it predicts when 1:4 (and perhaps 2:4 or 3:6 discriminations) will won't be possible for infants. Furthermore: It predicts that since ANS and subitizing representations compete for space in working memory, subitizing and ANS tasks can be pursued in tandem, provided that no more than three collections are considered at once. Thus, an infant will not discriminate a subitized collection of *one* from a subitized collection of *two* while prioritizing two approximately enumerated collections, but they might approximately enumerate a single large collection while simultaneously comparing two subitizable values.

There are philosophical upshots to my proposal too. I've argued that subitizing and ANS representations have a complex object-attribute structure, akin to an object file (Green & Quilty-Dunn 2021). My conjecture thereby builds upon recent philosophical work exploring the compositional structure of perceptual representations (Lande 2021) and representations with an analogue format (Clarke 2023; Lande 2024). What's distinctive is that, here, the objects picked out are *collections of items*, rather than bounded "middle-sized dry goods".

On the conjecture advanced, there needs to be an architectural, syntactic or semantic difference between the representations of collections picked out through subitizing and the representations of collections apt for approximate enumeration. This is because, on my account, the object-specifying element involved in a subitizing representation cannot have approximate number content freely attributed to it once the subitizing threshold is exceeded – something must prevent this, suggesting that the referential elements involved in subitizing and approximate enumeration will differ in kind (see Feigenson 2011).

No less importantly, however, the object-specifying referential elements in these representations, cannot be entirely devoid of content or significance for the infants who deploy them. Thus, unlike the sub-representational *fingers of instantiation,* that Pylyshyn and others posit to explain performance in MOT paradigms – symbols which are said to function as bare demonstratives, lacking content or accuracy conditions entirely (c.f. Echeverri 2017) – the collection-representing elements of subitizing representations enable infants to appreciate that it is *a collection* being picked out, which may or may not contain more items than a separate collection of one or two or three. For as we saw in Section 4.4, this is what prevents infants from reliably choosing subitizable collections (e.g., of 1) over non-subitizable collections (e.g., of 4) in Feigenson's tasks, and instead leaves them at chance. At the same time, it enables infants to systematically choose non-subitizable collections (e.g. of four) over empty collections (e.g., an empty bucket – see Section 4.4). So, beyond the fact that a complete account of perceptual structure may need to recognize that distinct symbols are employed when sustaining visual reference towards Spelke objects, subitizable collections and approximately enumerated collections (Feigenson 2011), my solution to the big-small problem suggests that these symbols need to be richer in content or semantic significance than purely demonstrative accounts of visual reference assume (e.g., Pylyshyn 2003).

(Main text, inc. footnotes and in-text references: 9,017 words)

Works cited:

1. Agrillo, C., & Bisazza, A. (2017). Understanding the origin of number sense: a review of fish studies. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *373*(1740), 20160511.
2. Beck, J. (2015). Analogue Magnitude Representations: A Philosophical Introduction. *The British Journal for the Philosophy of Science*, 66(4), 829–855. https://doi.org/10.1093/bjps/axu014
3. Beck, J. (2019). Perception is Analog: The Argument from Weber's Law. *The Journal of Philosophy*, 116(6), 319–349.
4. Beck, J. & Clarke, S. (forthcoming). Cardinality and autoscaling: Revisiting the content and format of the approximate number system. In Joonkoo Park, Eric Snyder & Richard Samuels, *Numerical Cognition: Debates and Disputes*.
5. Beran, M.J., Englund, M.D., Haseltine, E.L. *et al.* (2024). Monkeys overestimate connected arrays in a relative quantity task: A reverse connectedness illusion. *Atten Percept Psychophys* **86**, 2877–2887.
6. Brannon, E.M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition,* 83(3), 223-240.
7. Butterfill, S.A. and Sinigaglia, C. (2014), Intention and Motor Representation in Purposive Action. *Philosophy and Phenomenological Research*, 88: 119-145. https://doi.org/10.1111/j.1933-1592.2012.00604.x
8. Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS biology*, *4*(5), e125.
9. Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.
10. Cheyette, S.J. & Piantadosi, S.T. (2019). A primarily serial, foveal accumulator underlies approximate numerical estimation. *Proceedings of the National Academy of Sciences U S A.* 116(36):17729-17734.
11. Chiang, W.C. & Wynn, K. (2000). Infants' tracking of objects and collections. *Cognition* 77(3):169-95. doi:10.1016/s0010-0277(00)00091-3.
12. Clarke, S. (2022b). Beyond the icon: Core cognition and the bounds of perception. *Mind & Language*. https://doi.org/10.1111/mila.12315
13. Clarke, S. (2023). Compositionality and constituent structure in the analogue mind. *Philosophical Perspectives*. https://doi.org/10.1111/phpe.12182
14. Clarke, S. (2025). Number Nativism. *Philosophy and Phenomenological Research* 110(1): 226-52. https://doi.org/10.1111/phpr.13107
15. Clarke, S. & Beck, J. (2021). The number sense represents (rational) numbers. *Behavioral and Brain Sciences*, 44, E178. https://doi.org/10.1017/S0140525X21000571
16. Clarke, S., Qu, C., Luzzi, F. & Brannon, E., (2025). Children's number judgments are influenced by connectedness. *Developmental Science,* 28(4), e70032. https://doi.org/10.1111/desc.70032
17. Cordes, S., & Brannon, E. M. (2009a). Crossing the divide: Infants discriminate small from large numerosities. *Developmental Psychology, 45*(6), 1583–1594. https://doi.org/10.1037/a0015666
18. Cordes, S. and Brannon, E.M. (2009b), 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
19. Denison, S., & Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition*, 130(3), 335–347. *https://doi.org/10.1016/j.cognition.2013.12.001*.
20. Echeverri, S. (2017). Visual Reference and Iconic Content. *Philosophy of Science* 84 (4):761-781.
21. de Hevia, M.D., Izard, V., Coubart, A., Spelke, E.S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences* 111(13):4809-13. doi: 10.1073/pnas.1323628111.

22. Feigenson, L. (2011). Objects, sets, and ensembles. In S. Dehaene & E. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought* (pp. 13–22). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-385948-8.00002-5
23. Feigenson L, Carey S, Hauser M. (2002) The representations underlying infants' choice of more: object files versus analog magnitudes. *Psychological Science*. 13(2):150-156. https://doi.org/10.1111/1467-9280.00427..
24. Feigenson, L. & Halberda, J. (2004) Infants chunk object arrays into sets of individuals. *Cognition* 91(2): 173-190. https://doi.org/10.1016/j.cognition.2003.09.003
25. Feigenson L, Dehaene S, Spelke E. (2004) Core systems of number. *Trends in Cognitive Science* 8(7):307-314. https://doi.org/ 10.1016/j.tics.2004.05.002. PMID: 15242690.
26. Feigenson, L. and Carey, S. (2003), Tracking individuals via object-files: evidence from infants' manual search. *Developmental Science*, 6: 568-584. https://doi.org/10.1111/1467-7687.00313
27. Feigenson, L. and Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition* 97(3): 295-313. https://doi.org/10.1016/j.cognition.2004.09.010
28. Fodor, J.A. (1979). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences* 3(1): 63-73.
29. Fodor, Jerry A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.
30. Gallistel, C. R. (2011). Mental magnitudes. In S. Dehaene & E. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought* (pp. 3–12). Elsevier Academic Press.
31. Gallistel CR. (2017). The Coding Question. *Trends in Cognitive Sciences*. 21(7):498-508.
32. Gallistel, C. R. & Gelman, Rochel (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences*. 4(2):59-65. doi: 10.1016/s1364-6613(99)01424-2.
33. Gennari, G., Dehaene, S., Valera, C., & Dehaene-Lambertz, G. (2023). Spontaneous supra-modal encoding of number in the infant brain. *Current biology: 33*(10), 1906–1915.e6.
34. Giurfa M. (2019). An Insect's Sense of Number. *Trends in cognitive sciences*, *23*(9), 720–722.
35. Grech R, Cassar T, Muscat J, Camilleri KP, Fabri SG, & Zervakis M, et al. (2008). Review on solving the inverse problem in EEG source analysis. *Journal of Neuroengineering and Rehabilitation*. **5** (1): 25. doi:10.1186/1743-0003-5-25
36. Green, EJ. & Quilty-Dunn, 2020. What is an Object File? *British Journal for the Philosophy of Science*. axx055.
37. Halberda, J., Sires, S.F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science*. 17(7):572-6. doi: 10.1111/j.1467-9280.2006.01746.x.
38. Halberda, J. (2016). Epistemic Limitations and Precise Estimates in Analog Magnitude Representation. In *Oxford series in cognitive development. Core knowledge and conceptual change* (pp. 171–190).
39. Hyde, D.C. (2011). Two systems of non-symbolic numerical cognition. *Frontiers in Human Neuroscience*. 5:150. doi: 10.3389/fnhum.2011.00150.
40. Hyde, D.C., Boas, D.A., Blair, C., & Carey, S. (2010). Near-infrared spectroscopy shows right parietal specialization for number in pre-verbal infants. *Neuroimage*. 53(2):647-52. doi: 10.1016/j.neuroimage.2010.06.030.
41. Hyde, D.C. and Wood, J.N. (2011) Spatial attention determines the nature of nonverbal number representation. *Journal of Cognitive Neuroscience*. 23(9):2336-51. doi: 10.1162/jocn.2010.21581.
42. Intriligator, J. & Cavanagh, P. (2001). The spatial resolution of visual attention. Cognitive Psychology. 43(3):171-216. doi: 10.1006/cogp.2001.0755.
43. Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstractnumbers. *Proceedings of the National Academy of Sciences*, 106(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106
44. Jevons, W. S. (1871). The power of numerical discrimination. *Nature* 3, 363–372.

45. Kosslyn, S. (1980). *Image and Mind*. Cambridge: Harvard University Press.
46. Lande, K. J. (2021). Mental structures. *Noûs*, 55, 649–677. https://doi.org/10.1111/nous.12324
47. Lande, K. J. (2024). Pictorial Syntax. *Mind & Language,* 39: 518-539.
48. Libertus ME, Brannon EM. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science*. 13(6):900-906. https://doi.org/10.1111/j.1467-7687.2009.00948.x.
49. Libertus, M., Starr, A., & Brannon. E. (2014). Number trumps area for 7-month old infants. *Developmental Psychology* 50(1): 108-12.
50. Lipton, J.S. & Spelke, E.S. (2003). Origins of Number Sense: Large-Number Discrimination in Human Infants. *Psychological Science*, 14(5), 396–401.
51. Machery, E. (2014). In Defense of Reverse Inference. *British Journal for the Philosophy of Science* 65 (2):251-267.
52. Maley, C. J. (2011). Analog and digital, continuous and discrete. *Philosophical Studies*, 155(1), 117–131. https://doi.org/10.1007/s11098- 010- 9562- 8
53. Margolis, E. (2020). The Small Number System. *Philosophy of Science, 87(1):113-134. doi:10.1086/706087*
54. Marr, D. (1980). *Vision*. MIT Press.
55. 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. https://doi.org/10.1037/0097-7403.9.3.320
56. Moher, M. and Feigenson, L. (2011). Infants' abilities to parse and enumerate orthogonal ensembles. *Journal of Vision,* 11(11), 1284.
57. Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience,* 17(6): 366-382.
58. Pylyshyn, Z.W. (2003). *Seeing and Visualizing: It's not what you think.* MIT Press.
59. Pylyshyn, Z. (2007). Multiple Object Tracking. *Scholarpedia*, 2(10):3326.
60. Qu, C., Clarke, S., Luzzi, F & Brannon, E (2024). Rational Number Representation by the Approximate Number System. *Cognition* 250 (105839):1-13.
61. Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences* 46:e261.
62. Roitman, J.D., Brannon, E.M., & Platt, M.L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biology*;5(8):e208. doi: 10.1371/journal.pbio.0050208.
63. Scholl, B.J., & Leslie, A.M. (1999). Explaining the infant's object concept: Beyond the perception/cognition dichotomy. In E. Lepore & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 26–73). Oxford, England: Blackwell.
64. Spelke, E. S. (2003). What makes us smart? Core knowledge and natural language. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 277–311). Boston Review.
65. Spelke, E.S. (2022). *What Babies Know: Core Knowledge and Composition Volume 1.* Oxford University Press.
66. Srinivasan R (1999). Methods to Improve the Spatial Resolution of EEG. *International Journal of Bioelectromagnetism*. **1** (1): 102–11.
67. Starkey, P., & Cooper, R. G. (1980). Perception of Numbers by Human Infants. *Science*, *210*(4473), 1033–1035. http://www.jstor.org/stable/1684281
68. Starr, Ariel, Melissa E. Libertus, and Elizabeth M. Brannon. (2013) Infants show ratio-dependent number discrimination regardless of set size. *Infancy* 18.6: 927-941.
69. Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity pre-attentive stage in vision. *Psychological Review, 101*(1), 80–102. https://doi.org/10.1037/0033-295X.101.1.80
70. Wang, J. J., & Kibbe, M. M. (2024). "Catastrophic" set size limits on infants' capacity to represent objects: A systematic review and Bayesian meta-analysis. *Developmental science*, *27*(4), e13488.

71. Wood, J.N. & Spelke, E.M. (2005) Infants' enumeration of actions: numerical discrimination and its signature limits. *Developmental Science*. 8(2):173-81. doi: 10.1111/j.1467-7687.2005.00404.x.
72. Wynn, K. (1992). Addition and subtraction by human infants. *Nature* 358, 749–750. https://doi.org/10.1038/358749a0
73. Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science, 7*(3), 164–169. https://doi.org/10.1111/j.1467-9280.1996.tb00350.x
74. Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 74(1), B1–B11. https://doi.org/10.1016/S0010- 0277(99)00066- 9
75. Zosh, J. M., Halberda, J., & Feigenson, L. (2011). Memory for multiple visual ensembles in infancy. *Journal of Experimental Psychology: General*, 140(2), 141–158. https://doi.org/10.1037/a0022925