RESEARCH

Individuals and non-individuals in cognition and semantics: The mass/count distinction and quantity representation

Darko Odic[1], Paul Pietroski[2], Tim Hunter[3], Justin Halberda[4] and Jeffrey Lidz[2]

[1] University of British Columbia, CA

[2] University of Maryland College Park, US

[3] University of California Los Angeles, US

[4] Johns Hopkins University, US

Corresponding author: Jeffrey Lidz (jlidz@umd.edu)

Language is a sub-component of human cognition. One important, though often unattained goal for both cognitive scientists and linguists is to explicate how the meanings of words and sentences relate to the more general, non-linguistic, cognitive systems that are used to evaluate whether sentences are true or false. In the present paper, we explore one such relationship: an interface between the linguistic structures referring to individuals and non-individuals (specifically, count-nouns like *cows* and mass-nouns like *beef* in English) and the non-linguistic cognitive systems that quantify and compare number and area. While humans may be flexible in how they use language across contexts, in two experiments using standard psychophysical testing we find that participants evaluate a count-noun sentence (i.e., one including a pluralized noun, such as *blobs*) via numerical representations and evaluate a corresponding mass-noun sentence (i.e., one including a unmarked noun, such as *blob*) via non-numerical representations, consistent with a principled interface between language and cognition for evaluating these terms. This was the case even when the visual display was held constant across conditions and only the noun type was varied, further suggesting an important difference in how area and number, as well as count and mass nouns, are represented. These findings speak to issues concerning the semantics-cognition interface, the mass-count distinction, and the psychophysics of quantity representation.

Keywords: approximate number system; quantity representation; semantics-cognition interface; count/mass nouns; quantification

1 Introduction

A representational distinction between individuals (e.g., objects) and non-individuals (e.g., substances or extents) has played an important role in theories of cognitive representations (Scholl 2001; Feigenson 2007; Carey 2009) and in semantic theories focused on the formal structures that underlie linguistic meaning (Higginbotham 1994; Chierchia 1998b; Bale & Barner 2009; Rothstein 2010). For example, human infants and children have been shown to quantify and reason differently for individual objects than for piles of sand (Huntley-Fenner, Carey & Solimando 2002; Huntley-Fenner et al. 2002), suggesting that they represent objects as something more than mere aggregates of matter. Similarly, many human languages syntactically distinguish count nouns (e.g., *cow, chair*) from mass nouns (e.g., *beef, wood*), suggesting a difference in semantic representations.

At first blush, the count/mass distinction might seem to be a mere syntactic coding of the object/substance distinction. But this analogy is only apparent. Intuitively, mass

nouns like *jewelry* or *furniture* are used to refer to (collections of) individuals, as opposed to substances (Chierchia 1998a; Bale & Barner 2009). Count nouns like *line* or *twig* are used to talk about homogenous entities (in that any arbitrary subpart of a twig is a twig, just as an arbitrary subpart of water is water), further blurring a semantic distinction between count and mass (Mittwoch 1988; Krifka 1992). Indeed, the very things described with a plural count noun (e.g., *shoes, coins, ropes*) can often be described with a mass noun (e.g., *footwear, change, rope*), again indicating that the grammatical count-mass distinction does not align neatly with the psychological object-substance distinction. Finally, languages differ with regard to whether a lexical item is primarily a count or mass noun (e.g., *hair* is a mass noun in English, but a count noun in French; Chierchia 1998a).

Despite the lack of a clear reduction of count and mass nouns into representations of objects and substances, investigating the link between grammatical form and cognitive representation may nonetheless prove informative for two questions about human cognition. First, do basic intuitions of magnitude (e.g., how tall is a flagpole, how many people are in the room) reflect a single generalized magnitude system, or multiple such systems, each tuned to a specific dimension of experience (Walsh 2003; Cantlon, Platt & Brannon 2009; Lourenco & Longo 2010)? Second, is there a single kind of semantic representation from which we construct the meanings of both count and mass nouns (Chierchia 1998a; b), or are the meanings of count and mass nouns drawn from two similar but distinct representational domains (Link 1983; Landman 1991; Bale & Barner 2009)?

These two questions have more than a superficial similarity. In each case they ask whether apparently disparate representations are somehow unified, despite intuitive distinctions. They also ask about the unity and diversity of quantificational systems in linguistic and non-linguistic cognition. Surprisingly, psychological investigations of magnitude and linguistic investigations of quantification have largely proceeded without significant cross-fertilization. In the current paper, we will argue that the parallels between linguistic and nonlinguistic quantification can be leveraged to inform theorizing in both domains through the interface between linguistic expressions and the cognitive systems used to verify their meanings.

Moreover, in order to understand how children acquire the count/mass distinction in language, it is important to first have a clear understanding of how this distinction is represented in linguistic semantics and how it relates to cognitive magnitude representations. The latter is especially important, as these cognitive representations also undergo some development that spans the time when relevant linguistic representations are acquired (Halberda & Feigenson 2008; Odic et al. 2013).

The simple act of assessing whether "More of the dots are blue than yellow" in Figure 1 requires engaging a broad array of cognitive systems. To understand this sentence, the reader must identify the meanings of the individual words and, by using basic rules of syntactic organization and semantic composition, determine how meanings of words combine in this sentence to form a larger unified meaning. But to verify the sentence – i.e., to evaluate it, as understood, for truth or falsity – one must invoke psychological capacities like visual attention, numerical magnitudes, and ordinal comparison, each of which behaves according to its own rules that are distinct from those of natural language. Language *use* thus depends on the existence of a tractable interface between our linguistic-semantic representations and our psychology.[1] And, of course, learning the meanings of words like *more* will depend, at least to some extent, on these same interfaces.

The difficulty of characterizing this interface has been a major stumbling block in integrating theories of quantification in linguistics and cognitive psychology, and more

---

[1] Readers interested in the interface between language and cognition will find further details concerning the Interface Transparency Thesis, an idea about how linguistic algorithms and cognitive algorithms may interact(Pietroski et al. 2009; Lidz et al. 2011).
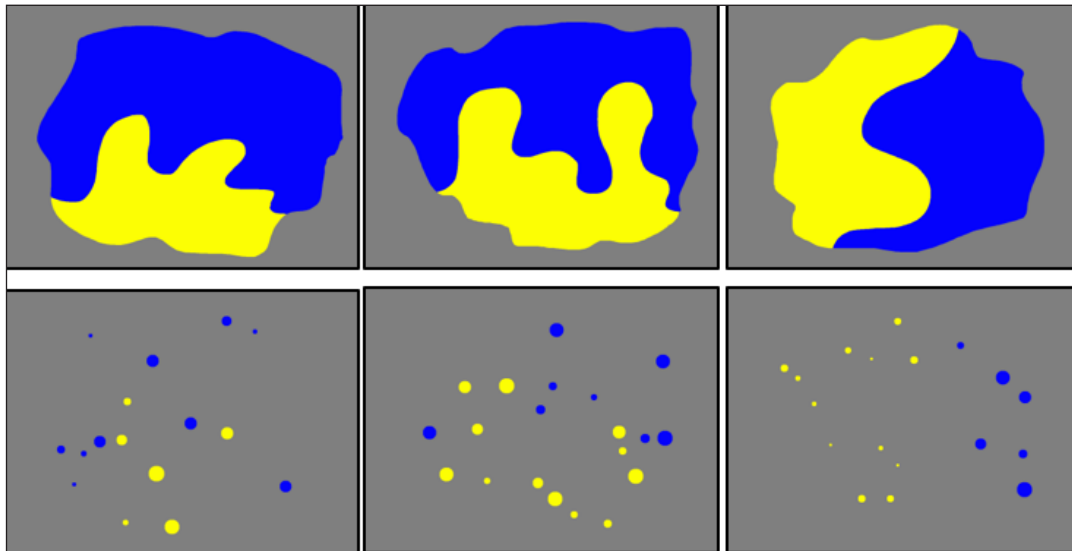
Figure 1: Example of stimuli used in Experiment 1.

generally in rigorously integrating linguistic semantics with the rest of cognition (Pietroski et al. 2009; Lidz et al. 2011). As noted above, an important open question in semantics is whether count and mass terms rely on formally distinct semantic representations (Link 1983; Landman 1991; Bale & Barner 2009) or a common underlying formal structure (Chierchia 1998a; b); and an important open question in psychology is whether conceptions of number and area rely on distinct cognitive systems (Castelli, Glaser & Butterworth 2006; Cohen Kadosh, Lammertyn & Izard 2008) or a single unified magnitude system (Walsh 2003; Bueti & Walsh 2009; Lourenco & Longo 2010). By exploring the interface between lexical semantics and magnitude representations, we will argue that there are multiple distinct cognitive systems for quantification and that the count nouns and mass nouns link up to distinct semantic representations.

In the first experiment, we investigate what behavioral signatures, if any, may differentiate number (object) processing from area (substance/extent) processing with stimuli that are either quite clearly about number (sets of dots) or about area (a single continuous mass). Then, in the second experiment, we turn to the question of whether the semantic distinction of count and mass nouns interfaces with the cognitive representations of number and area, even given identical displays. The combined results indicate that there are two distinct cognitive systems at play in quantification – one for quantifying number, the other for quantifying area – and that the linguistic count/mass distinction connects with these cognitive systems in way that certain classes of semantic theories would not predict.

It is now well accepted among psychologists that humans can represent number in at least two ways. The first method, and the one probably most familiar to us all, is by counting and representing number exactly (Gelman & Gallistel 1978; Wynn 1992; Feigenson, Dehaene & Spelke 2004; Carey 2009). However, although such a representational system is useful, it only emerges after a lot of learning (Gelman & Gallistel 1978; Wynn 1992; Carey 2009), and it may also require a spoken/signed language (Gordon 2004; Frank et al. 2008; Carey 2009). An alternative number representational system – the Approximate Number System (ANS) – appears to be innate in both humans and other animals (Dehaene 1997; Feigenson et al. 2004; Izard et al. 2009) and is used by infants (Feigenson et al. 2004) and adults who lack number words (Pica et al. 2004) to make numerical discriminations and

compute the outcomes of addition and subtraction events.[2] The ANS is what gives us an intuitive feel of how many things are in a set, such as, for example, in guessing how many marbles are in a jar. In the present experiments concerning number, we focus on this gut intuitive sense of numerosity generated by the ANS.

The ANS is not capable of representing number exactly. Instead, it approximates number, and represents it as a continuous Gaussian activation (for details, see Results) of several numerical values on a mental number line (Dehaene 1997; Nieder & Miller 2004). Thus, one never has knowledge of exactly how many items are in a scene – merely a rough range. Additionally, the comparison of two such activations is successful insofar as the two representations do not overlap too much – the greater the degree of overlap between two approximate number representations, the more difficult it is to discriminate between them (Green & Swets 1966). This property of the ANS results in its compliance to Weber's law – the smaller the ratio of two numbers, the worse discrimination is between them, regardless of the total number of items. Thus, for relatively high ratios, like 2.0 (10 blue: 5 yellow dots), discrimination is easy, while for relatively low ratios, like 1.2 (12 blue: 10 yellow dots), discrimination is hard and error-prone. Compliance of a numerical judgment performance to Weber's law is the primary behavioral signature of ANS use.

Individuals vary in how well their ANS can discriminate numbers. An individual's discrimination abilities are measured by the internal precision of the representation – or the Weber fraction (w; Green & Swets 1966). The Weber fraction roughly corresponds to the most difficult ratio that an observer can discriminate with 75% accuracy and indicates the amount of "noise" in the internal representations (the Gaussian distributions) that make up the dimension. A person with a higher Weber fraction for a given dimension will have noisier internal representations and have a harder time discriminating between two representations within the dimension (e.g., some people will easily discriminate 10 from 8 dots, while others struggle with this discrimination). These individual differences are well behaved and can be estimated for each person by precise mathematical models.

The ANS supports a sense of numerical magnitude. But humans and non-human animals can represent other magnitudes as well. These other magnitude representations also rely on a noisy, Gaussian representational format (Feigenson 2007; Cantlon et al. 2009). Decades of work on various cognitive continua, including length, brightness, pitch, time, and area have suggested that humans represent various "approximate" dimensions, which all follow Weber's law (e.g., 10 seconds versus 5 seconds is easier to discriminate than 12 seconds versus 10 seconds).

This similarity in discrimination behavior (i.e., discrimination that obeys Weber's law) has led several researchers to propose that a single, domain general magnitude system may underlie all our judgments about quantity (Walsh 2003; Bueti & Walsh 2009; Lourenco & Longo 2010). Under one version of this view, our quantity representations do not differentiate between objects and extents – both object-related quantities, like number, and extent-related quantities, like area, are encoded on the exact same mental quantity line by identical sets of noisy Gaussian representations. In Experiment 1, we put this idea to the test, and look for differences in the Weber fraction that describes the underlying noise signature for area and number discrimination within a subject.

While discrimination in many dimensions (e.g., number, area, brightness, loudness) obeys Weber's law, these dimensions may not all have precisely the same Weber fraction

---

[2] A separate system for representing number, often termed the Parallel Individuation System, is thought to underlie precise representations of 2–4 objects, but lack set-based cardinality operations (Feigenson et al. 2004). In order to avoid the possibility of participants using the parallel individuation system, all stimuli in our experiments have at least 5 items, forcing the use of the ANS.

within an individual (Feigenson 2007). It is possible that, for example, area information may be represented with higher precision (i.e., a lower Weber fraction) than number information, and this difference would be consistent with different internal processes being engaged when representing and verifying the values of these dimensions. If there were only a single domain general magnitude system for both area and number, then, at a first pass, the Weber fraction for representing area should be the same as that for representing number (though perhaps small systematic differences might arise from low-level perceptual differences in processing each type of information). Even then, however, because a common representation is giving rise to all magnitude discriminations, the Weber fractions should at least be strongly, if not perfectly, correlated, as they are measuring the exact same parameter.

On the other hand, if number discrimination performance results in different, uncorrelated, Weber fractions (e.g., area is better than number), then two distinct magnitude systems may be involved, each tuned to just one of these two dimensions (e.g., an Approximate Number System, and a separate Approximate Area System). A distinction between area and number processing can then be further explored in the interface with language.[3] In the present experiments we assess the Weber fraction for number and for area and compare these.

Previous research on area Weber fractions has been very mixed, with some work suggesting that area representations show poor Weber fractions (Morgan 2005) and others suggesting that it shows relatively good Weber fractions (Nachmias 2008), and correlations between area and number tasks have never been examined. Likewise, some work has suggested that, given a choice of encoding a set of objects by either number or area, number tends to be preferred by children (Gathercole 1985; Barner & Snedeker 2006; Cantlon, Safford & Brannon 2010) and infants (Cordes & Brannon 2008). Due to the conflicting literature, it is unclear that we have evidence for or against a similarity in Weber fractions between number and area representations.

To investigate the precision of area discrimination, in Experiment 1 we presented adult subjects with non-geometric figures and asked them to discriminate which of two colors was larger in area (i.e., "Is more of this blob blue or yellow"; Figure 1a). These images were also created so that the competing dimensions of number, line length, and aspect-ratio could not be used (Morgan 2005; Nachmias 2008). This was contrasted with a task where the same images were converted into displays of dots where total area was varied, and subjects had to answer a count-noun question which, given the stimuli presented, clearly required number discrimination (i.e., "Are more of these dots blue or yellow"; Figure 1b). Performance across ratios was modeled with a psychometric equation to determine the Weber fraction that best describes approximate number and approximate area discrimination. Should the two Weber fractions differ, we would have some initial evidence for a different behavioral patterns in area and number processing that may be indicative of a difference in the representational systems used. Additionally, should the two judgments use different representations, we should expect no correlations in Weber fractions across the two tasks. We first test for different Weber fractions for area and number processing using unambiguous displays in Experiment 1 (e.g., Figure 1), and then turn to displays that are ambiguous between area and number in Experiment 2.

---

[3] For psychologists interested primarily in the single magnitude system versus multiple systems debate, data concerning Weber fractions is certainly relevant and has yet to be explored within subjects, but this exploration is not the primary focus for the present article. Rather, we hope to show how any differences in number and area processing can impact theorizing in other fields (e.g., formal semantics) and how looking at the interface between cognition and language opens up new sources of evidence for both psychologists and linguists.

2 Experiment 1
2.1 Methods
2.1.1 Subjects

Participants were 16 college-age adults, naïve to the purpose of the experiment, who either volunteered or were compensated $10 for their time.

2.1.2 Materials & apparatus

Each participant was individually tested in a dimly lit room. The experiment was presented on a Macintosh Pro with a 22" LCD screen. Participants were seated about 42 cm away from the monitor with their heads unrestrained. All programs were custom made in a Java environment.

During the Area task, participants were presented with a blob image that was divided into a yellow and blue part (Figure 1a). The generation of blob images was done in three steps. In the first step, we generated 26 unique outlines; care was taken to ensure that the outlines were curvilinear and natural, resembling a mass of stuff on a page. In the second stage, the outlines were divided into a blue and yellow area. For each outline, we generated 132 splits of blue and yellow. The lines that cut the blob into two areas were also made to look natural and curvilinear. This method gave us a broad range of areas and perimeters for the blue and yellow sections, while retaining a non-geometric look (cf. Tegthsoonian 1965). In the final stage, the blue and yellow areas were measured using a custom-made pixel-counting program. Ratios were determined by dividing the larger area (in number of pixels) by the smaller area. Overall, we generated over 3,400 blob images, but only administered those with ratios varying from the easy 2.2 (approx. 11:5 pixels) to the very difficult 1.01 (approx. 101: 100 pixels); the displays with ratios of over 2.2 were deemed too easy during pilot testing, and were not administered.

For the Number task, participants were presented with an image of blue and yellow dots on the screen (Figure 1b). To create the dot images, we took our blob images and used a custom-made Python program to extract circles of various radii from the blob areas. During the extraction, one of three area parameters were used to determine the size relationship between the blue and yellow dots: dots were either correlated in area and number (with the winning color being larger by the same ratio in both area and number), anti-correlated (with the same ratio in both but giving opposite answers; e.g., blue wins in number but yellow wins in area), or were area controlled (area was matched in two dot sets). The dot-images had four fixed ratios: 2.0 (12:6 or 14:7 dots), 1.67 (10:6 or 15:9 dots), 1.5 (9:6 or 12:8 dots) or 1.2 (12:10 or 18:15 dots).

2.1.3 Procedure

Participants were seated in front of the monitor and were instructed on their task. Each participant did both the Area and Number task, with order counterbalanced across subjects. During the Area task, participants were asked to indicate whether "More of the blob is 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" (note that *blob*, like *rock* and *string*, is flexible between a mass and count reading, but that the context of the sentence and the image clearly implies that it is used as a mass noun).[4] In the Number task, subjects were asked to indicate whether

---

[4] In order to further alleviate this issue, we ran a new group of 20 participants in an additional Experiment. These participants completed the Area task with the same stimuli as in Experiment 1 and we instructed half (N = 10) to verify whether "More of the blob is blue or yellow", and half (N = 10) to verify whether "More of the goo is blue or yellow". *Goo*, unlike *blob*, is unambiguous in English and favors a mass interpretation. We found no effect of which sentence was used suggesting that the all participants relied on a mass interpretation of *blob* in our tasks (i.e., the word *blob* in singular, much like *rock* and *string* is understood as a mass noun).

"More of the dots are blue or yellow", and to press the F key for "More of the dots are yellow", and the J key for "More of the dots are blue".

Each trial began with a number in the center of the gray screen that indicated the remaining number of trials. Participants had to press the spacebar to begin the trial, and were told that, if tired, they could take as long as they needed before starting each individual trial. After the spacebar was pressed, the stimuli appeared for 500 milliseconds (ms), and were backward masked by an image that had several dozen small blue and yellow blobs (the same masking image was used in both tasks). No image appeared more than once. In both tasks, there were 10 practice trials at the beginning, identical to the test trials, which were excluded from analysis. In the Area task, there were 300 trials; in the Number task, where there were three types of area-controlled displays, there were 600 trials, which were evenly distributed across ratios and area-controls.

After the experiment was over, the participants were debriefed. On average, the experiment took 30 minutes to complete.

2.2 Results

Our analysis was done in two parts. In the first part, we examine whether the two tasks demonstrated compliance to Weber's law. In the second part, we model which Weber fraction best describes the performance on each task, and, subsequently, compare the performance on the two tasks.

In order to determine whether the Area task showed an effect of ratio and to maximize statistical power, we rounded and binned the continuously distributed ratios into six evenly spaced ratio bins. Because these bins were not identical to the Number task, we ran two separate Analyses of Variance (ANOVAs) rather than an single, omnibus one. Assuming that performance will comply with Weber's law, the two tasks can be directly compared by comparing the resulting Weber fractions, a content-independent metric of the noise in the underlying representations; this analysis is presented further below.

In the case of the Area task, we ran a 2 (Order: Area-First, Number-First) $\times$ 6 (Ratio: 1.1, 1.3, 1.5, 1.7, 1.9, 2.1) Mixed-Measures ANOVA that showed a significant effect of Ratio ($F(5,70) = 114.12$; $p < 0.01$) and accounted for 89% of the variance; there was no effect of Order ($F(1,14) < 1$) or an Order $\times$ Ratio interaction ($F(5,70) < 1$). For the Number task, we ran a 2 (Order: Area-First, Number-First) $\times$ 4 (Ratio: 1.2, 1.5, 1.67, 2.0) Mixed-Measures ANOVA that also showed an effect of Ratio ($F(3,42) = 160.313$; $p < 0.01$) and accounted for 92% of the variance; once again, there was no effect of Order ($F(1,14) = 2.764$; $p > 0.10$) nor an interaction ($F(3,42) = 2.313$; $p > 0.10$). Therefore, in both cases, there was a ratio-dependent effect that is consistent with Weber's law (Figure 2).

Next, the two conditions were modeled to determine the Weber fractions. The model used to describe the performance is one that is used widely in the psychophysics literature (Green & Swets 1966; Pica et al. 2004), where n_1 and n_2 refer to the quantity of each set (e.g., 20 and 10 dots), w refers to the Weber fraction, and $erfc$ refers to the complimentary error function of a Normal/Gaussian curve:

$$\frac{1}{2} erfc \left(\frac{n_1 - n_2}{\sqrt{2}w \sqrt{n_1^2 + n_2^2}} \right) *100$$

Extensive details on this model are presented in the Appendix to Lidz et al., (2011), and are only described briefly here. The model assumes that the underlying representations are distributed along a continuum of Gaussian/Normally distributed random variables.
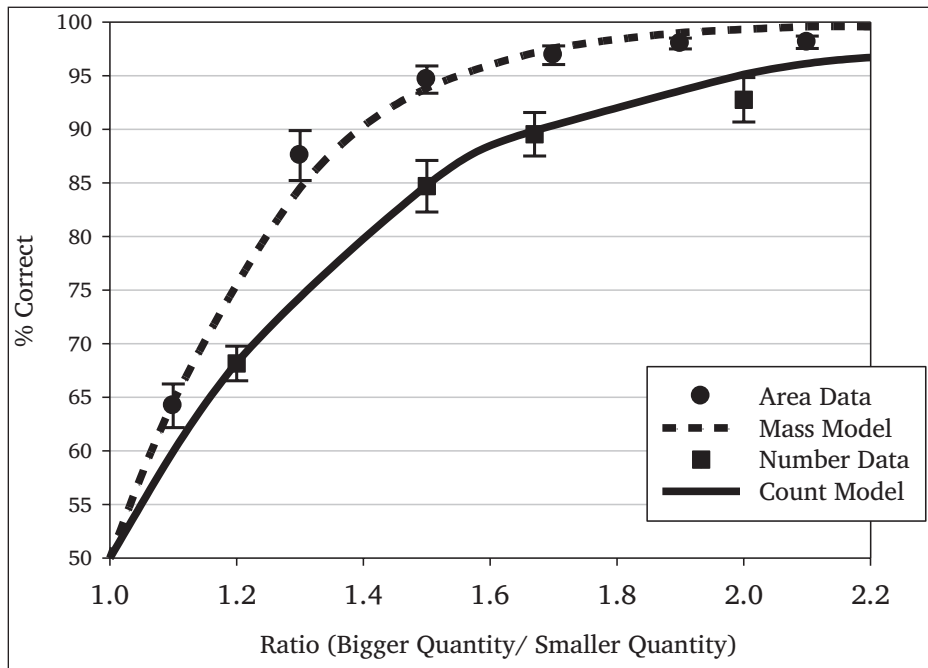
Figure 2: Data from Experiment 1.

Because each representation (e.g., one triggered in response to 20 dots) is distributed across the continuum, two overlapping values, be they two numbers or sizes of blobs, will naturally representationally overlap, creating confusion. In other words, as the ratio of two quantities becomes increasingly similar (i.e., closer to a ratio of 1.0), their Gaussian representations should overlap more and participants should have a more difficult time determining which is larger resulting in decreasing accuracy at the task as a function of ratio. If both number and area processing comply with Weber's law, this same basic model can be used to fit subjects' performance with the resulting Weber fraction indicating the amount of noise in the underlying Gaussian representations of number and area.

This model has only a single free parameter – the Weber fraction (w) – which indicates the amount of noise in the underlying Gaussian representations (i.e., the standard deviation of the Gaussian number or area representations). Larger w values indicate poorer discrimination of the system across all ratios. The best fitting w value was determined for each subject using the least-squares method, minimizing the squared error between the model and each observed data point. The modeled group data are presented in Figure 2.

Both the Number and Area tasks were well-described by the Gaussian psychophysical model (both $r^2 > 0.97$ for the group fits, Figure 2), confirming that Weber's law applied and returning an estimate of the Weber fraction for each task. Next, we examined whether a single Approximate System underlies performance in both tasks (as would be revealed by a non-significant difference in Weber fraction between tasks) or whether there are two distinct Gaussian systems, the Approximate Number System (ANS) and an "Approximate Area System" (AAS), which would be revealed by a significant difference in Weber fraction between these two tasks.

Performance from each subject for each task was fit independently and the w values were compared. The average w value for the Area task was 0.18 (*comparable with estimate for area perception in* (Morgan 2005); Standard Error/SE = 0.02), while the average w for the Number task was 0.27 (*comparable with estimate for number perception in* (Izard & Dehaene 2008); SE = 0.03). These values were run through a paired-sample t-test which

showed a significantly lower Area w (t(15) = –3.534; p < 0.01). We also examined, participant-by-participant, which w value was lower – for all but one participant the w value for the Area task was lower than the Number task.

This result suggests that two different approximate systems – an Approximate Area System (AAS) and an Approximate Number System (ANS) – were engaged on the two tasks and that the AAS has less noise than the ANS across participants. One concern, however, is that this difference in Weber fraction may be due to a single magnitude system being used to make discriminations on different types of perceptual evidence. One way to address this is by considering the individual differences in Weber fraction across the two tasks. If each person relied on a single system (e.g., the ANS) in the two tasks then we would expect individual performance on the two tasks to correlate. However, the Weber fractions on the two tasks did not correlate (p > 0.25) suggesting independent sources of representational noise and, thus, that two different approximate systems were being used on the two tasks.

Another test of the independence of number and area processing is to determine if the area-control manipulation within the Number task had any effect on performance. Number trials were split into the three area-control conditions used to create the displays (i.e., area-correlated, area-anti-correlated, area-controlled), and a Weber fraction (w) for each subject for each condition was determined via least squares. The average w for area-controlled trials was 0.28 (SE = 0.03), for area anti-correlated was 0.32 (SE = 0.06) and for area-correlated was 0.24 (SE = 0.01). A 3-level (Condition: Area-Correlated, Area-Anti-Correlated, Area-Controlled) Repeated Measures ANOVA found no effect of condition (F(2,30) = 2.134; p > 0.13) suggesting that area correlations did not impact participants' number decisions. This remained true even when the area-controlled condition was removed and we compared only the two most extreme trial types (i.e., area-correlated and area-anti-correlated), suggesting that area content was not used in estimating number (Hurewitz, Gelman & Schnitzer 2006; Barth 2008). The results of Experiment 1 suggest that area and number discrimination engaged distinct magnitude systems or, at the very least, distinct representations of the display.

2.3 Discussion

In our first experiment, we found that number discrimination and area discrimination are each consistent with Weber's law. We also found a significant within-subject difference in the Weber fraction estimated from these two tasks suggesting that number and area estimation rely on distinct cognitive systems (i.e., an Approximate Number System – ANS – and an Approximate Area System – AAS). The possibility of distinct cognitive systems for number and area was further supported by the lack of a correlation between the Weber fractions estimated for these two tasks and the absence of an effect of area-control on the number estimation trials.

A potential concern may be that participant's Weber fractions differed because of some inherent difference in the display – one display may have been easier to quantify and compare than the other. Note that if this was the case, we would expect a correlation between the number and area tasks, which we did not find; likewise, we found no influence of area on number. Ideally, however, we should expect to find a distinction between number and area processing in *identical* stimuli. We turn to this question, as well as the mapping of count and mass nouns to number and area processing, in the second experiment.

Given that we have some preliminary evidence about the distinction between number and area processing in cognition, we can now turn to the problem of how the linguistic count/mass distinction interfaces with general cognition and makes contact with the

cognitive distinction between objects and substances/extents.[5] We return to the discussion of multiple quantity systems in the general discussion.

3 The mass/count distinction

The mass/count distinction has been studied extensively, and we will not attempt a review here (for representative discussions see Link 1983; Higginbotham 1994; Chierchia 1998a; b; Bale & Barner 2009; Rothstein 2010). There is disagreement about how to characterize the distinction in a theoretically illuminating way,[6] but, for present purposes, two standard syntactic diagnostics will suffice, at least for languages such as English:

a) Only count nouns can be pluralized (cow/cows, beef/*beefs); relatedly, only count nouns can combine with numerical determiners (three cows, *three beef/beefs).

b) Certain determiners only combine with count nouns (many dots/*many mud/muds); others only combine with mass nouns (much mud/*much dot/dots); and some, of particular interest here, can combine with either kind of noun (more dots/more mud).

Any diagnostics for the count/mass distinction must come with the caveat that there is considerable flexibility in how nouns can be used (e.g., Frisson & Fraizer 2005). Even paradigmatic count nouns like *dinosaur* can have odd-sounding mass noun counterparts, as in "After the meteor struck, there was dinosaur all over the place", and paradigmatic mass nouns like *mud* can have odd-sounding count noun counterparts, as in "At the spa, we tried three different muds." And many nouns seem perfectly comfortable in either mode, as in "The blue rocks and guitar strings were found on some blue rock and old string". This leaves it open whether the homophony is due to multiple lexical entries that are semantically related (Frisson & Fraizer 2005) or multiple derivations from a common lexical root.

Several theories have been proposed to account for the linguistic data concerning the mass/count distinction. Barner and Snedeker (2006) and Bale and Barner (2009) argue that count nouns always refer to individuals, and, given comparative count-noun sentences, are verified via number. Mass nouns, under this account, usually refer to non-individuals, and are not necessarily verified via number. One prediction of this theory is that the differences in the truth conditions between mass/unmarked and count/pluralized noun comparative sentences like "More of the blob is blue" and "More of the blobs are blue" will give rise to different verification procedures. Specifically, a quantification system that represents number should be used for count-noun sentences and a different quantification system that represents area, volume, or brightness should be used for mass-noun sentences.

Another prominent mass/count theory has been put forth by Chierchia (1998a; b) and argues that both count nouns and mass nouns refer to individuals or units, but that

---

[5] We wish to highlight again that we are not seeking to reduce the semantic count/mass distinction to a psychological object/substance distinction. While the semantic distinction is not reducible, count terms or mass terms may find a preferred mode of verification in the psychological magnitude representations of number and area (or continuous extents like mass).

[6] For example, many count/mass nouns exhibit (so-called) atomicity/homogeneity: intuitively, a cow can divide into smaller individuals, but no such sub-individual is a cow. At the same time, any portion of beef is also beef, even though beef does not divide naturally into beef atoms. But this is not a definitive criterion: consider the mass nouns *furniture* and *succotash* (Bale & Barner 2009; Rothstein 2010) argues that the count noun *fence* is also a counterexample. And as noted above, the difference is not ontological. Languages also often differ with regard to whether a lexical item (e.g., *hair/chevaux*) is primarily as a mass or count noun (Chierchia 1998b).

these units are vague for mass nouns and thus need to be identified during verification (cf., Rothstein 2010). Under at least some readings of this view, the verification procedures for mass noun and count noun sentences invoke one and the same non-linguistic quantification system, namely the one that discriminates number: the speaker needs to identify the relevant unit of the visual image referred to by the mass noun, and must then count up the units. Thus, given two buckets of paint, one may judge which one has more paint by deciding that the unit of paint is a small 1×1 inch square, and then counting up the squares in each bucket (we return to discussing Chierchia's 2015 account in more detail in the General Discussion).

Chierchia (1998a; b) and Rothstein (2010) posit only a weak distinction between count and mass nouns in the semantics while, Link (1983), Landman (1991) and Bale & Barner (2009) suggest there is a stronger formal distinction. Evidence from number and area cognition may be relevant to this debate, but the distinction between number and area processing demonstrated in Experiment 1 does not necessitate that there is a strong count/mass distinction in the semantics of the sort that Bale & Barner (2009) propose. For example, perhaps participants in the Area task who heard the mass noun *blob* were biased towards verifying the sentence via a numerical quantification systems, but, given that there was only a single blob, no numerical information was available, and participants opted for the next best thing – area discrimination (Rothstein 2010).

A stronger test of the existence of two independent magnitude systems in cognition (i.e., ANS and AAS) and of the interface between these systems and a prominent count/mass distinction in the semantics would be to use identical displays and *only* vary the question asked (i.e., by varying the minimal syntactic difference between count and mass nouns). If count and mass nouns differ in their reference, we should find significant differences in the Weber fraction estimates for these two conditions. In particular, if count syntax maps to number processing and mass syntax to area processing, we should find Weber fractions comparable to those found in the Experiment 1.
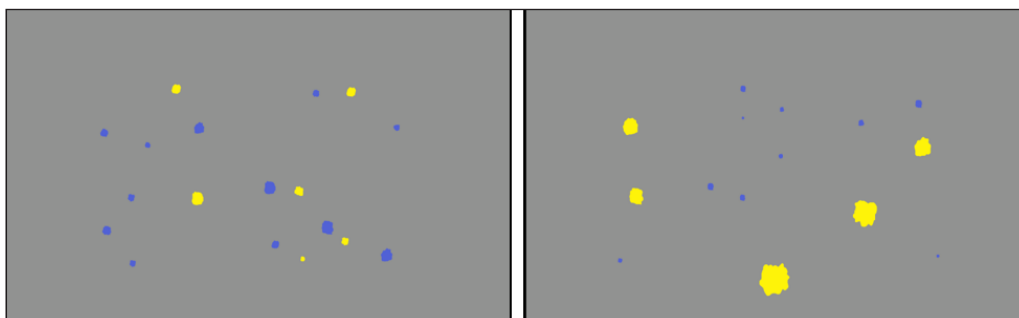
4 Experiment 2
4.1 Methods

4.1.1 Subjects
Participants were 12 adults, naïve to the purpose of the experiment, who either volunteered or were compensated $10 for their time. None had participated in the first experiment.

4.1.2 Materials and apparatus
Every factor from the first experiment was retained except for the following. Participants were presented with a display containing several blue and yellow colored blobs (Figure 3). The blobs were randomly selected from a set of 18 curvilinear outlines and randomly

Figure 3: Example of stimuli used in Experiment 2.

placed on the screen. We used five ratios for both Mass and Count comparisons: 2.0, 1.5, 1.2, 1.14, and 1.12.

On half of the trials, the total summed area of the colored blobs was correlated with the number (e.g., blue wins by both *more dots* and *more area)*, and on the other half of the trials, the total summed area of the colored blobs was anti-correlated with number (e.g., blue wins by *more dots* but yellow wins by *more area*); area-controlled trials were removed as these trials would not generate an answer for the area question. Importantly, as in Experiment 1 the ratio by number and by area was identical on each trial, but inverted in the anti-correlated condition (e.g., if the number of dots was in a ratio of 2:1 with more yellow, then the number of pixels was in a ratio of 2:1 with more blue). This ensured that subjects saw stochastically identical displays for the count noun and mass noun conditions.

4.1.3 Procedure

Each participant did both the Count and the Mass task, with order counterbalanced across subjects. In order to further minimize any differences between the two tasks and have a consistent sentence structure, we used the noun *blob* in both instances, varying only the count/mass syntax through the use of either the singular/unmarked form of the noun, or the plural form. Thus, during the Mass 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 the Count 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". Thus, all participants saw identical displays and pushed identical buttons and only the *is/are* and *blob/s* changed in the initial instructions. Our question was whether this small change in syntax would result in subjects recruiting different verification procedures and, thus, distinct non-linguistic magnitude systems as revealed by different Weber fractions. We predicted that success at the Mass task (i.e., "Is more of the blob blue or yellow") would engage the Approximate Area System (AAS) while success at the Count task (i.e., "Are more of the blobs blue or yellow") would engage the Approximate Number System (ANS) and this difference would be reflected in a difference in Weber fraction even when the displays were stochastically identical.

Our display time remained at 500 ms, but stimuli were not masked (pilot testing reveled no effect of mask and so it was removed as some subjects found it distracting). There were 300 trials per condition. After the experiment was over, the participants were debriefed. On average, the experiment took 30 minutes to complete.

4.2 Results

We followed the same analyses we performed in Experiment 1. We first ran a 2 (Order: Mass-First, Count-First) × 2 (Task: Mass, Count) × 5 (Ratio: 2.0, 1.5, 1.2, 1.14, and 1.12) Mixed Measures ANOVA on percent correct at each ratio. There were no main effects or interactions of any factor with Order (all $p > 0.20$), suggesting that pragmatic effects of contrast between the two tasks are not responsible for our results. There was a significant effect of Ratio ($F(1,40) = 140.66$; $p < 0.01$) and of Task ($F(1,10) = 10.93$; $p < 0.01$). Consistent with Experiment 1, participants did significantly better in the Mass condition (Mean = 0.80; SE = 0.01) than the Count condition (Mean = 0.74; SE = 0.02).

Next, we turned to modeling. The modeled group performance is presented in Figure 4. Just like in Experiment 1, the Weber fraction was fitted for each participant for each condition. The average *w* for the Mass condition was 0.20 (SE = 0.05; group fit $r^2 = 0.99$) and for the Count condition was 0.29 (SE = 0.13; group fit $r^2 = 0.98$); reflecting better discrimination in the Mass condition. This difference was significant as measured by a
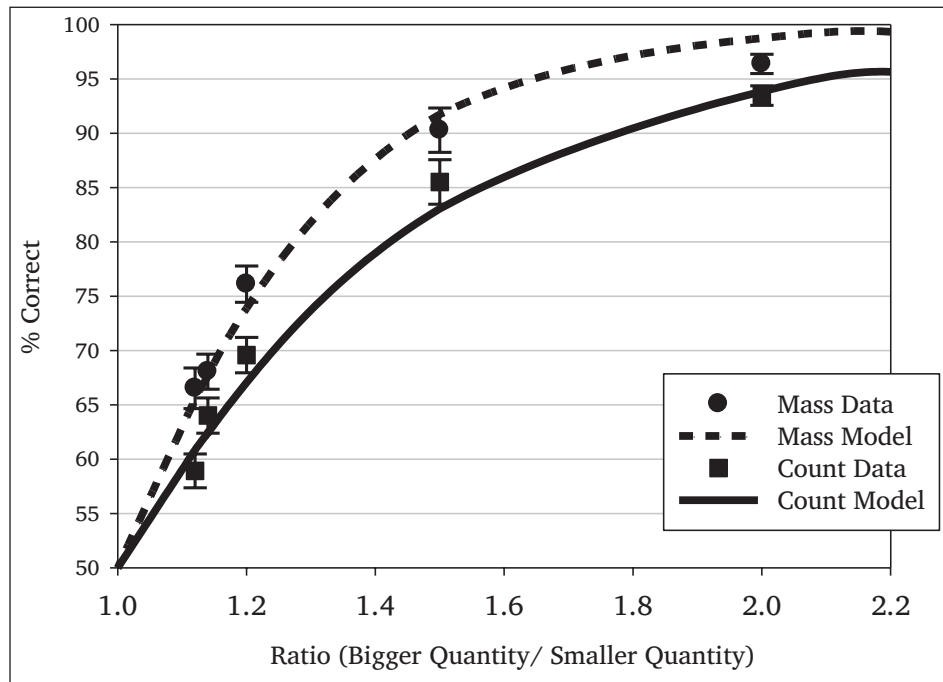
Figure 4: Data from Experiment 2.

planned t-test ($t(11) = -2.428$; $p < 0.05$). Both these values closely mirror the values found in Experiment 1 for the Area and Number tasks. We also examined, participant-by-participant, which w value was lower – for all but one participant, the w value for the Mass task was lower than the Count task.

Thus, it may be possible that the same magnitude system was used (e.g., the ANS), but that blob-area units are somehow easier to verify than blob-number units. If this were the case, the Weber fractions should be correlated across the two conditions. However, as in Experiment 1, this correlation was not significant ($p > 0.30$). These results provide no evidence that the same magnitude system was used in both tasks.

As a final assessment of the independence of mass and count, we separated the trials into those where area and number correlated and those where they did not. In the case of the Count task, there was no difference between these two stimulus array types ($t(11) = -1.78$; $p > 0.10$), replicating the finding from Experiment 1. In the case of the Mass task, there was a significant difference ($t(11) = -2.428$; $p < 0.05$), with the anti-correlated trials (i.e., where number gives the opposite answer of area) being superior (Mean = 0.16; SE = 0.06) to the correlated trials (Mean = 0.26; SE = 0.08).

Two explanations are possible for this difference in the Mass condition. First, processing number in some way interfered with processing area. Although this proposal is possible, it seems unlikely given that participants were *worse* on those trials where number and area *agreed* on an answer. A second explanation seems more likely: when number and area are *anti*-correlated the color that wins in area has much larger blobs than the color that wins in number (e.g., 5 yellow blobs that are twice as big as 10 blue blobs). Thus, there are two methods of finding the answer: either by summing and comparing the total area, or by comparing the largest blob in each set (in the correlated condition, only the former strategy is possible). Either one of these strategies is consistent with the participants using area rather than number, but the anti-correlated trials allow for an additional source of evidence (i.e., largest blob), thus allowing for better discrimination performance. Consistent with this suggestion, data from the Count task in Experiment 2 suggest that when participants were judging number, differences in area were efficiently

ignored as there was no difference in performance between correlated and anti-correlated trials in this task.

5 General discussion

In Experiment 1 we found that both number and area discrimination obey Weber's law but that these tasks result in distinct and significantly different Weber fraction (i.e., area discrimination is better than number discrimination). In Experiment 2 we found that this distinction between number and area processing is maintained when subjects are asked to make number and area judgments about identical displays. These results demonstrate that number and area processing are distinct and engage separate magnitude representations (i.e., the Approximate Number System – ANS – and the Approximate Area System – AAS).

Our data is incompatible with cognitive theories that claim that numeric and non-numeric quantity representation are largely identical (Walsh 2003; Lourenco & Longo 2010). If this is the case, it is unclear why, when given identical displays, two different Weber fractions capture the participant's performance. Clearly, there must be *some* difference in what the participants do when the sentence meaning suggests that they should gather and compare numeric information than when the sentence meaning suggests that they should gather area information. Although differences in encoding the stimuli may be responsible for a difference in Weber fractions, this seems especially unlikely given the identical displays in Experiment 2 and given the lack of correlations between the two tasks in both experiments. Therefore, some difference in the internal noise of *independent* quantity systems, is likely responsible for this difference in Weber fractions (for evidence of a separation of brain regions that process area and number, see Castelli et al. 2006; Cohen Kadosh et al. 2008)

We also explored the interface between the linguistic representation of mass and count syntax and the psychological representations of area and number. The results of Experiment 2 suggest that when verifying a comparative sentence containing mass noun syntax, participants are biased towards a cognitive system whose acuity is different from the cognitive system that they are biased towards when verifying a minimally different comparative sentence containing count noun syntax. Specifically, count noun syntax appears to bias towards numerical quantity as the relevant quantity and are, therefore, verified by the ANS (or, given sufficient time, counting), and mass noun syntax (given our stimuli) specified area quantity, and are, therefore, verified by the AAS. Although the present work only directly speaks to *more*, other determiners, such as *most*, should demonstrate equivalent results.

Note that we are not claiming that all mass noun verifications need to occur via the AAS, nor that all count noun verifications need to occur via the ANS: e.g., given sufficient time, participants may have chosen to count the items, and given mass nouns like *furniture* participants may have used the ANS. Our claim is, instead, that comparative count noun sentences bias towards processing *number* through whatever cognitive system can represent it given the demands of the task, and that mass nouns bias towards whatever the relevant type of quantity given by the noun or context is (see below for details).

Our results stand in contrast to and inform several theories about the count/mass distinction in English, about individual and non-individual representation and comparison, and about the semantics-cognition interface.

First, one might suggest – in keeping with a long tradition in semantics – that details of how truth conditions should be specified are largely independent of how linguistic expressions happen to interface with particular cognitive systems; cp. Davidson (1974). One might imagine genuinely holistic minds such that specific word or sentence meanings

do not constrain which cognitive systems are used to compute an answer to the question posed. For such thinkers, the ANS might be employed more often with count nouns because this tends to makes evaluation easier, or because we are familiar with enumerating whole objects (which count nouns typically refer to), and not because of the semantic content of count nouns. Thus, under such a view, there would be no relationship between truth conditions and verification procedures. Instead, this view leads us to expect only effects of the display's interaction with many possible cognitive systems.

But if language had no influence over the selection of relevant systems for verification, then given the same stimuli, the same cognitive system should have been used for verification. Looking back at the second experiment, however, it is clear that the count noun question resulted in much poorer performance than the mass noun question, despite the visual stimuli being identical. The sole factor that differed between these two conditions was the linguistic form used, and so, at least prima facie, the linguistic form and the specific cognitive systems used to verify it are intimately linked.

Our data also present a challenge for any semantic theory that treats both mass nouns and count nouns as having meanings that are specified (perhaps vaguely) in terms of countable atomic individuals. Other things equal, such theories predict that participants will either use whatever cognitive system is most easily accessed (which, given the above paragraph, cannot be sustained), or that there is a bias for identical verification procedures regardless of the presence of count or mass nouns. In particular, given a view that both count and mass nouns are ultimately unit-based, one should expect that the ANS verification procedure will be used in both instances. The difference, of course, would be that, in the case of area, one would need to assign a unit of area (e.g., a 10×10 pixel box), while, in the case of most count-nouns (e.g., cows, dots, people), units are immediately available.

Several things suggest that this is not what subjects did. First, if number was the quantity computed in both mass and count cases, whatever representational and processing noise is affecting one task should affect the other. However, in our own data, we found no correlation between the two tasks, suggesting two distinct cognitive systems were being used. Second, if there is some form of cognitive conversion from area into number, then additional noise should exist in the case of area as a product of this conversion, resulting in a higher Weber fraction when compared to number. However, our data suggests exactly the opposite – subjects were, across the board, better at computing area. Finally, independent work from (Castelli et al. 2006) demonstrates a difference in how the brain computes area and number in displays similar to those from our Experiment 1. There appears no reason to think, then, that countable individuals – the default unit of the ANS system – are anything like the "units" of area used by the AAS.

Our data is thus in line with those who maintain that the syntactic distinction between count and mass terms has a correlated semantic distinction; see (Barner & Snedeker 2006; Bale & Barner 2009), who suggest that the correct way to separate mass and count nouns is by what type of quantity they seek out during verification. The idea is that count nouns have a feature that biases towards numerical comparisons while mass nouns lack any such feature and require context to determine what approximate system it should use for verification (including, in the case of furniture, perhaps the ANS itself). Our data are consistent with this view.

An important issue concerns languages that do not have a count/mass distinction, such as Cantonese, Mandarin, and Japanese, which instead communicate about discrete units using classifier phrases, analogous to the English *bottles of water*. At first glance, classifier languages may behave differently from the patterns observed here: given that a classifier transparently provides a unit by which the object should be subdivided, one may make a

reasonable prediction that the ANS will be invoked for most classifier phrases. Recently, our group empirically tested this prediction by recruiting native Cantonese speakers, and giving them the exact same stimuli from Experiment 2, varying only whether the classifier phrase referred specifically to number, or more openly referred to portions of blue things. Contrary to the above prediction, we found that Cantonese speakers verified identical stimuli using two distinct systems – while the numeric classifier led to verification by the ANS, the portion classifier led to verification by the AAS (Odic et al. in prep.). Hence, the patterns observed in the two experiments reported here may even transparently extend to classifier languages.

In a series of more recent publications, Chierchia (2010; 2015) has updated and elaborated on his proposal that the count/mass distinction is rooted in a more basic distinction between substances and objects. In this work, Chierchia claims that the meanings of mass nouns like *water* are akin to the meanings of plural nouns like *cats*, with an important twist: portions of water, unlike pluralities of cats, exhibit a special kind of divisional *vagueness* that precludes natural counting of minimal parts. The leading idea is that *water* represents examples of water as being (for practical purposes) endlessly divisible into examples of the same kind, whereas *cats* represents examples of (many) cats as being divisible into examples of the same kind only if the division respects the unity of individual cats. Chierchia's updated proposal is sufficiently programmatic to avoid predictions about how the contrast between pluralized noun and mass nouns should map to strategies for evaluating sentences in our test situations. But, although space prohibits us from discussing it at length, we wish to note that several key features of Chierchia's proposal seem puzzling given our data; see also Pietroski (forthcoming).

Chierchia's (2010; 2015) proposal centers around the idea that mass nouns, such as *water*, are relevantly like pluralized nouns, such as *cats*. The idea is that plural nouns apply to distinctive objects – collections of some kind – which have countable elements. Thus, although there may be some vagueness (e.g., given evolution or bizarre cases) about what exactly counts as a cat, once such vagueness is resolved there is no further question about how many cats there are in a set. By contrast, there are many equally good (but overlapping) ways of carving a typical sample of water into multiple samples. On his view, *water* applies to certain collections, each portion of which is also water, but these collections fail to be enumerable in the relevant sense. Two problems emerge from this suggestion. First, rather than abandoning the apparent link between mass nouns and substances, Chierchia is committed to treating mass nouns such as *furniture* and *jewelry* as "fake mass nouns", despite ample evidence that, for example, children acquire these nouns at the same time as any other "real" mass noun (Barner & Snedeker 2008). Second, Chierchia's proposal invokes several problems with division and molecules: isn't a single molecule of H_2O (an example of) water in any possible world, and won't this account require that each such molecule be constituted by submolecular water particles? Chierchia's reply is to relativize his technical formulation to "natural contexts", characterized as sets of worlds that are shared by competent (but typically scientifically naïve) speakers (e.g., perceiving the smallest quantity of water without any specialized machinery). But, we don't understand how linguistic competence can presuppose scientific ignorance, and apparently preclude the very hypothesis of atomism or the absence of complex machinery (where presumably, eyes aided by contact lenses are not complex, but microscopes are).

Chierchia seems to be saying that in contrast with nouns like cat, nouns like water have remarkable meanings that somehow carry substantive (and false) implications about perceivable quantities of stuff. Perhaps this suggestion will turn out to be correct. One could then accommodate our experimental findings – in which the context provides a salient notion of unit that is neither minimal nor vague – by saying that the grammatical

property of being a mass noun triggers a measuring strategy that is appropriate for substances *across* contexts – because substances exhibit divisional vagueness – while the grammatical property of being a count noun requires a counting strategy. But we suspect that all things considered, a more plausible package of views will combine our findings with the idea that mass (i.e., non-count) nouns have neutral meanings that allow for a measuring strategy, while the more complex count nouns have more restrictive meanings that call for counting.

Finally, the work here broadly illustrates the usefulness of studying the interface of semantic and cognitive theories (Pietroski et al. 2009; Lidz et al. 2011). Our results suggest that, when the scene is simple and the cognitive systems involved are well-understood (e.g., the ANS and AAS), there is a lawful interface between the semantics involved in understanding a sentence and the psychology involved in verifying if the sentence is true; and in this way, empirical work can inform both psychological theorizing and semantic theorizing.

In our previous work, we have suggested the Interface Transparency Thesis (ITT), which claims that the meaning of sentences exerts a bias in verification towards cognitive systems that most naturally implement the operations expressed in the meaning of the sentence. Thus, given that a sentence "More of the dots are blue" includes a request for a comparison operation (via the word *more*), those cognitive systems that have the ability to compute comparisons will be biased towards (i.e., most likely to be used, all else being equal) during verification. The present work also demonstrates that the operand – the noun – provides further bias, as the presence of the count noun in the above construction also biases towards those cognitive systems that can compare *and* represent number (e.g., ANS, or, given sufficient time, verbal counting). In the case of mass nouns this bias is especially striking, since alternative verification procedures, including ones related to ANS and individuating the blobs, were available.

Thus, our results demonstrate an important similarity in both how we should treat individuals and non-individuals, for the purposes of quantification and comparison, in the semantics and in cognition. This tight relationship between semantic distinctions and cognitive distinctions should not be a surprise to anyone; in fact, it is necessary for meaning and verification to successfully occur every time we use language. What is more surprising, then, is that there has been such a large divide between cognitive and formal semantic theories of meaning. Through this work, we hope to highlight both how semantic and cognitive theories can mutually aid one another. Theories of semantics based in the truth-conditional properties of expressions can provide justification and predictions for theories of how these properties are mentally represented. And, we can draw on methods of assessing mental representations and processes from cognitive science in order to distinguish semantic theories that make similar (or identical) predictions with respect to the strictly linguistic properties of expressions. By enriching our conception of linguistic meaning to include more than representation-independent truth conditions, and by having cognitive theories constrained by the formal work of semantics, we hope to ultimately provide a theory of semantics that is both cognitively and linguistically justified.

Abbreviations

AAS = Approximate area system, ANOVA = Analysis of variance, ANS = Approximate number system, *erfc* = Gauss error function, SE = standard error, *w* = Weber fraction.

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

References

Bale, Alan C. & David Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26(3). 217–252. DOI: https://doi.org/10.1093/jos/ffp003

Barner, David & Jesse Snedeker. 2006. Children's early understanding of mass-count syntax: Individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development* 2(3). 163–194. DOI: https://doi.org/10.1207/s15473341lld0203_2

Barth, Hilary. 2008. Judgments of discrete and continuous quantity: An illusory Stroop effect. *Cognition* 109(2). 251–266. DOI: https://doi.org/10.1016/j.cognition.2008.09.002

Bueti, Domenica & Vincent Walsh. 2009. The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1525). 1831–1840. DOI: https://doi.org/10.1098/rstb.2009.0028

Cantlon, Jessica F., Kelley E. Safford & Elizabeth M. Brannon. 2010. Spontaneous analog number representations in 3-year-old children. *Developmental Science* 13(2). 289–297. DOI: https://doi.org/10.1111/j.1467-7687.2009.00887.x

Cantlon, Jessica F., Michael Platt & Elizabeth M. Brannon. 2009. Beyond the number domain. *Trends in Cognitive Sciences* 13(2). 83–91. DOI: https://doi.org/10.1016/j.tics.2008.11.007

Carey, Susan. 2009. *The origin of concepts*. New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780195367638.001.0001

Castelli, Fulvia, Daniel E. Glaser & Brian Butterworth. 2006. Discrete and analogue quantity processing in the parietal lobe: A functional MRI study. *Proceedings of the National Academy of Sciences* 103(12). 4693–4698. DOI: https://doi.org/10.1073/pnas.0600444103

Chierchia, Gennaro. 1998a. Plurality of mass nouns and the notion of 'semantic parameter'. *Events and Grammar* 70. 53–103. DOI: https://doi.org/10.1007/978-94-011-3969-4_4

Chierchia, Gennaro. 1998b. Reference to kinds across language. *Natural Language Semantics* 6(4). 339–405. DOI: https://doi.org/10.1023/A:1008324218506

Chierchia, Gennaro. 2015. How universal is the mass/count distinction? Three grammars of counting. In Yen-hui Audrey Li, Andrew Simpson & Wei-Tien Dylan Tsai (eds.), *Chinese syntax: A cross-linguistic perspective*, 147–177. Oxford, England: Oxford University Press.

Cohen Kadosh, Roi, Jan Lammertyn & Veronique Izard. 2008. Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology* 84(2). 132–147. DOI: https://doi.org/10.1016/j.pneurobio.2007.11.001

Cordes, Sara & Elizabeth M. Brannon. 2008. The difficulties of representing continuous extent in infancy: Using number is just easier. *Child Development* 79(2). 476–489. DOI: https://doi.org/10.1111/j.1467-8624.2007.01137.x

Davidson, Donald. 1974. Belief and the basis of meaning. *Synthese* 27(3). 309–323. DOI: https://doi.org/10.1007/BF00484597

Dehaene, Stanislas. 1997. *The number sense*. New York, NY: Oxford University Press.

Feigenson, Lisa. 2007. The equality of quantity. *Trends in Cognitive Sciences* 11(5). 185–187. DOI: https://doi.org/10.1016/j.tics.2007.01.006

Feigenson, Lisa, Stanislas Dehaene & Elizabeth S. Spelke. 2004. Core systems of number. *Trends in Cognitive Sciences* 8(7). 307–314. DOI: https://doi.org/10.1016/j.tics.2004.05.002

Frank, Michael C., Daniel L. Everett, Evelina Fedorenko & Edward Gibson. 2008. Number as a cognitive technology: Evidence from Piraha language and cognition. *Cognition* 108(3). 819–824. DOI: https://doi.org/10.1016/j.cognition.2008.04.007

Frisson, Steven & Lyn Frazier. 2005. Carving up word meaning: Portioning and grinding. *Journal of Memory and Language* 53(2). 277–291. DOI: https://doi.org/10.1016/j.jml.2005.03.004

Gathercole, Virginia C. 1985. More and more and more about "More". *Journal of Experimental Child Psychology* 40. 73–104. DOI: https://doi.org/10.1016/0022-0965(85)90066-9

Gelman, Rochel & Charles Randy Gallistel. 1978. *The child's understanding of number*. Boston, MA: Harvard University Press.

Gordon, Peter. 2004. Numerical cognition without words: Evidence from Amazonia. *Science* 306(5695). 496. DOI: https://doi.org/10.1126/science.1094492

Green, David Marvin & John A. Swets. 1966. *Signal detection theory and psychophysics*. Newport Beach, CA: Peninsula Pub.

Higginbotham, Jim. 1994. Mass and count quantifiers. *Linguistics and Philosophy* 17(5). 447–480. DOI: https://doi.org/10.1007/BF00985831

Huntley-Fenner, Gavin, Susan Carey & Andrea Solimando. 2002. Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition* 85(3). 203–221. DOI: https://doi.org/10.1016/S0010-0277(02)00088-4

Hurewitz, Felicia W., Rochel Gelman & Brian Schnitzer. 2006. Sometimes area counts more than number. *Proceedings of the National Academy of Sciences* 103(51). 19599–19604. DOI: https://doi.org/10.1073/pnas.0609485103

Izard, Veronique, Coralie Sann, Elizabeth S. Spelke & Arlette Streri. 2009. Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences* 106(25). 10382–10385. DOI: https://doi.org/10.1073/pnas.0812142106

Izard, Veronique & Stanislas Dehaene. 2008. Calibrating the mental number line. *Cognition* 106(3). 1221–1247. DOI: https://doi.org/10.1016/j.cognition.2007.06.004

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Sag Ivan & Szabolcsi Anna (eds.), *Lexical matters*, 29–53. Stanford University, CA: CSLI Publications

Landman, Fred. 1991. *Structures for semantics*. New York, NY: Springer. DOI: https://doi.org/10.1007/978-94-011-3212-1

Lidz, Jeffrey, Justin Halberda, Paul Pietroski & Tim Hunter. 2011. Interface transparency thesis and the psychosemantics of most. *Natural Language Semantics* 19(3). 227–256.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Meaning, Use and Interpretation of Language* 21. 302–323. DOI: https://doi.org/10.1515/9783110852820.302

Lourenco, Stella F. & Matthew R. Longo. 2010. General magnitude representation in human infants. *Psychological Science* 21(6). 873–881. DOI: https://doi.org/10.1177/0956797610370158

Mittwoch, Anita. 1988. Aspects of English aspect: On the interaction of perfect, progressive and durational phrases. *Linguistics and Philosophy* 11(2). 203–254. DOI: https://doi.org/10.1007/BF00632461

Morgan, Michael J. 2005. The visual computation of 2-D area by human observers. *Vision Research* 45(19). 2564–2570. DOI: https://doi.org/10.1016/j.visres.2005.04.004

Nachmias, Jacob. 2008. Judging spatial properties of simple figures. *Vision Research* 48(11). 1290–1296. DOI: https://doi.org/10.1016/j.visres.2008.02.024

Nieder, Andreas & Earl K. Miller. 2004. A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences of the United States of America* 101(19). 7457–7462. DOI: https://doi.org/10.1073/pnas.0402239101

Odic, Darko, Melissa E. Libertus, Lisa Feigenson & Justin Halberda. 2013. Developmental change in the acuity of approximate number and area representations. *Developmental Psychology* 49(6). 1103–1112. DOI: https://doi.org/10.1037/a0029472

Pica, Pierre, Cathy Lemer, Veronique Izard & Stanislas Dehaene. 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306(5695). 499–503. DOI: https://doi.org/10.1126/science.1102085

Pietroski, Paul, Jeffrey Lidz, Tim Hunter & Justin Halberda. 2009. The Meaning of 'Most': Semantics, numerosity and psychology. *Mind & Language* 24(5). 554–585. DOI: https://doi.org/10.1111/j.1468-0017.2009.01374.x

Rothstein, Susan. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27(3). 343–397. DOI: https://doi.org/10.1093/jos/ffq007

Scholl, Brian J. 2001. Objects and attention: The state of the art. *Cognition* 80(1–2). 1–46. DOI: https://doi.org/10.1016/S0010-0277(00)00152-9

Tegthsoonian, Martha. 1965. The judgement of size. *The American Journal of Psychology* 78(3). 392–402. DOI: https://doi.org/10.2307/1420573

Walsh, Vincent. 2003. A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences* 7(11). 483–488. DOI: https://doi.org/10.1016/j.tics.2003.09.002

Wynn, Karen. 1992. Children's acquisition of the number words and the counting system. *Cognitive Psychology* 24(2). 220–251. DOI: https://doi.org/10.1016/0010-0285(92)90008-P