Disambiguating quantity judgements: mass/count and extra-grammatical cues

Sven Smeman, Maaike Smit, James A. Hampton, & Yoad Winter*

Abstract. Comparative quantity judgements are a useful probe into the semantics of the mass/count distinction, where count nouns usually trigger cardinal comparisons (*more dogs*), and mass nouns trigger non-cardinal measurement (*more rice*). However, exceptions like 'object' mass nouns (*furniture*) and 'mixed' comparatives (*more gold than diamonds*) complicate this pattern. In such cases there is often a mismatch between the mass/count status of the noun and the criterion for comparison, which challenges our understanding of the mass/count distinction and how it affects quantity judgements. We propose that these mismatches reflect a systematic ambiguity, where the mass/count distinction is one of the factors influencing disambiguation. Using a new experimental method focused on ambiguity judgements instead of truth-value judgements, the results support the traditional semantic encoding of the mass/count distinction, with operations of 'packaging' and 'grinding' triggered by extra-grammatical factors.

Keywords. mass/count nouns; comparatives; disambiguation; grinding; packaging

1. Introduction. The mass/count distinction in English is often viewed as a grammatical reflection of semantic discreteness, with count nouns (e.g. *cats*) typically referring to discrete objects and mass nouns (e.g. *water, justice*) representing substances or abstract concepts without a clear atomic structure. However, English and many other languages have 'object' nouns like *furniture* and *jewelry* that challenge this correlation, as their referents are discrete despite being grammatically mass. Recent analyses suggest weaker links between grammatical countability and semantic discreteness, either by downplaying the role of grammar with object mass nouns (Gillon 1992) or proposing discrete interpretations of such nouns (Barner & Snedeker 2005, Chierchia 2010). Further research has shown that quantity judgements also rely on pragmatic pressures and the perceived discreteness of real-world referents (Scontras et al. 2017, Rothstein 2017), which are not always clear-cut, leaving the semantics of the mass/count distinction an ongoing area of debate.

This paper studies potential ambiguities in quantity judgements with comparatives. Instead of focusing on categorical judgments of cardinality or measurement as in Barner & Snedeker's work, we explored participants' perception of ambiguities between these strategies. Results show that ambiguity in comparatives is closely tied to the mass/count distinction, both with 'flexible' nouns (e.g. *more rocks*/*rock*) and with 'object' mass nouns (e.g. *more bags*/*baggage*). However, simple noun expressions (*bags, baggage, rock*/*s*) do not support the same levels of unambiguous counting or measurement that is seen with expressions like *number of bags*/*rocks* or *volume of baggage*/*rock*. These results reflect the gradient of perceived ambiguity shown in Figure 1.

The ambiguity of quantity judgments has important implications for theories linking the mass/count distinction to continuous/discrete meanings. While our results show systematic evidence

*Work on this paper was supported by a grant of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 742204). Sven Smeman, Maaike Smit, & Yoad Winter: Utrecht University, James A. Hampton: City St. George's, University of London.

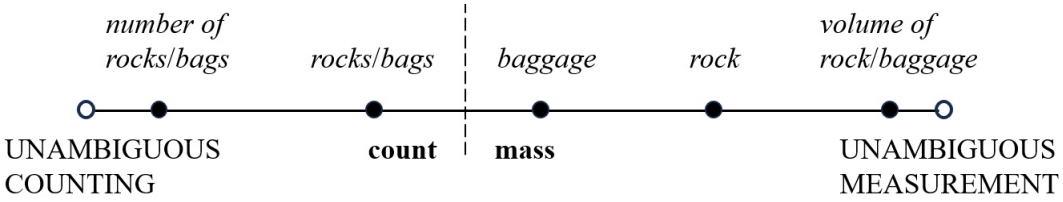

Figure 1: Scale of ambiguity in quantity judgements

for this connection, it is just one factor influencing counting and measurement. The behavior of 'object' mass nouns, previously seen as exceptional, aligns with their conventionally discrete referents. However, even 'substance' mass nouns like *beer* can exhibit counting effects in contexts that highlight packaging, though less so than object mass nouns. Likewise, the strong tendency for count nouns to trigger cardinal comparisons can be contextually inhibited. We argue that these observations support a theory that applies 'grinding' and 'packaging' operations based on pragmatic triggering. In this theory, grinding and packaging are not only triggered by mismatches between syntax and lexical semantics as in *too much rabbit* ('grinding' of a count noun) or *three beers* ('packaging' of a mass noun), but also without such mismatches as in *more rabbits* (or *more beer*), in case pragmatics triggers grinding (packaging, resp.), leading to exceptional quantity judgements.

2. Background: mass/count distinctions and quantity comparisons. In examples like (1), count nouns (CNs, e.g. *bag/s*) are in the plural and support a cardinal comparison:

(1) Anna has more bags/clay/baggage than Ben.

By contrast, mass nouns (MNs) are in the singular and show an irregular behavior: 'substance' MNs (SMNs, e.g. *clay*) prefer non-cardinal measurement (by volume, weight, etc.) while 'object' MNs (OMNs, e.g. *baggage, furniture*) trigger cardinal comparisons (McCawley 1975). Barner & Snedeker (2005) experimentally studied quantity judgments by presenting participants with comparative questions on visual stimuli as in Figure 2. The stimuli depicted two quantities, one of which with fewer discrete items whose total size is larger. This creates a scenario where counting and measurement should yield different answers to the question *who has more* ⟨noun⟩*?*. The OMNs in this study all showed the same (near unanimous) levels of counting as the CNs.

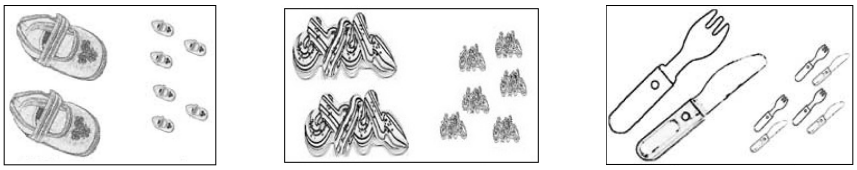

Figure 2: *who has more shoes/toothpaste/silverware?* (Barner & Snedeker 2005)

Figure 3: *who has more?* (Scontras et al. 2017)

The special properties of OMNs have led researchers to relax the strict correspondence between the mass/count distinction and continuity/discreteness of noun meanings. Gillon (1992) suggested that while all CNs have discrete, countable meanings, the grammar of MNs is 'mute' on discreteness. Thus, MNs are expected to support discrete/continuous interpretations only on the basis of extra-grammatical factors (e.g. world knowledge). This line does not offer any general-

ization on the semantics of MNs. By contrast, Barner & Snedeker introduce a sharp distinction between two types of MNs: SMNs vs. OMNs. OMNs are likened to CNs using a syntactic feature (+IND = 'individual') that distinguishes both noun classes from SMNs (−IND).[1] A problem for this account comes from examples like (2) below (Rothstein 2017, Grimm & Levin 2012):

(2) John has more (#pieces of) furniture than Bill, so he should use the larger moving truck.

Rothstein argues that OMNs like *furniture* in (2) can be measured by volume, unlike CNs like *tables*, which must be counted. Rothstein generalizes that CNs are counted in comparatives, SMNs are measured, and OMNs can be either counted or measured.[2] This view aligns with Gillon's, attributing the SMN/OMN distinction to extra-grammatical factors. Scontras et al. (2017) experimentally showed the importance of such factors by observing that non-cardinal measurement often occurs with objects perceived as substances, even if not explicitly referred to as MNs (Figure 3). Hampton & Winter (2024) further support Rothstein's claims experimentally, demonstrating a greater reliance on non-cardinal measurement for OMNs like *baggage* compared to CNs like *bags* in contexts favoring such comparisons.

3. Ambiguity in quantity comparisons. What is the general role that the mass/count distinction and extra-grammatical factors play in the quantity interpretation of nouns? This question involves two empirical problems:

> ***CN problem***: To what extent can CNs compromise count-based interpretations?
>
> ***MN problem***: To what extent can MNs compromise measure-based interpretations? Specifically, do OMNs support counting as strongly as CNs?

There is general agreement that OMNs frequently allow cardinality-based interpretations in comparatives despite their syntactic 'mass' type. However, following Hampton & Winter (2024), we question whether OMNs reject measure-based interpretations to the same extent as CNs. This prompts further investigation into 'exceptional' quantity interpretations of both MNs and CNs, as illustrated in the following examples:

(3) Anna ate more beans/peas/lentils than Ben.

(4) Mary put two (and a half) oranges in the punch.

(5) a. Pirates' treasures usually contained more gold than diamonds.
b. He had more hair than teeth, and his hairs totalled three.

In (3), we are more likely to refer to amounts of beans, peas and lentils than to their cardinalities (McCawley 1975). Sentence (4) predominantly reports on the cardinality of oranges, but at least for some speakers, it also involves a 'container' reading, which refers to amounts of orange juice (Snyder 2021). Comparatives that 'mix' CNs and MNs as in (5) may lead to a non-cardinal measurement of *diamonds* as in (5-a), and cardinality judgements about *hair* as in (5-b) (Winter

[1] Chierchia (2010) makes a similar distinction in semantic terms: SMNs have an 'unstable' atomic structure whereas OMNs have atoms that are as 'stable' as those of CNs. Thus, like Barner & Snedeker, Chierchia assumes a categorical distinction between nouns in terms of their 'discreteness', putting OMNs and CNs in one class and SMNs in another.

[2] For Rothstein, CNs grammatically require counting, whereas OMNs support counting through measurement using numerosity estimation.

2022). These cases support the idea that 'mismatches' between a count/mass status and a cardinality/measurement interpretation are not limited to OMNs.

Mismatches between *syntactic* noun classification and lexical mass/count preferences are well-documented. For example, a typical CN like *bicycle* can appear as syntactically 'mass' (*there's bicycle all over the floor*), while *beer*, usually an MN, can appear as countable (*three beers*). Building on these phenomena, we propose that CNs (MNs) without syntactic mismatch can also invoke non-cardinal (cardinal, resp.) quantity judgments, driven by extra-grammatical or pragmatic factors. Thus, phrases like *a lot of bicycles* (*beer*) or *more bicycles* (*beer*) may involve non-cardinal (cardinal, resp.) quantity judgments, making OMN behavior less exceptional than thought.

By weakening the semantic connection between count/mass and cardinal/non-cardinal quantity judgements, we do not intend to deny the strong effect of the mass/count distinction on interpretation. Indeed, following Hampton & Winter (2024), we hypothesize that CNs are *more easily* interpreted using cardinal interpretations than MNs, and conversely: it is easier to apply non-cardinal measurement to MNs than to CNs. To summarize, we present two hypotheses:

(H_1) Discreteness is semantically encoded in CNs, whereas MNs encode continuity.
(H_2) Comparatives grammatically allow measurement and counting with both CNs and MNs, where the choice is affected by the mass/count distinction, together with pragmatic factors and the perception of real-world objects.

To test these hypotheses we experimented with quantity interpretations, and examined to what extent they are affected by the mass/count distinction and by grammar-external factors. In our experiments, as in previous experimental work, we asked speakers to contribute linguistic judgements on comparative statements relative to situations where counting and measurement lead to different results. For example, in one of our tests we presented participants with sentences like *A has more bags/baggage than B* in situations where in terms of cardinality A has fewer bags, but A's bags occupy a larger volume than B's bags. Unlike Barner & Snedeker (2005), we did not use forced choice questions about a single comparative sentence. Rather, participants were requested to simultaneously report their judgements about two comparative sentences, of the forms:

(6) a. *A has more* ⟨noun⟩ *than B* b. *B has more* ⟨noun⟩ *than A*

We asked participants to indicate whether they accept both sentences in (6) or only one of them. If speakers can employ both counting and measurement, they are expected to choose the first option; if they only employ one strategy, they are expected to choose only one of the sentences.

This method follows recent works that highlight problems in evaluating semantic theories using simple truth-value judgements (Syrett & Musolino 2015, Pinto & Zuckerman 2019). As Syrett & Musolino point out, speakers may have robust biases against one of the readings of a given sentence although their semantics does not disallow it. When sentences like (6-a) and (6-b) are presented in isolation, truth-value judgements may put a pressure on speakers to reject one of these sentences just because their salient strategy makes the other sentence true. For instance, in relation to Figure 2 speakers may reject a sentence like *A has more shoes than B* because they strongly prefer the cardinality-based interpretation of CNs, hence *B has more shoes than A* would be far preferred as their description of the situation. Presenting both sentences simultaneously aims to reduce the effect of this potential confound.

packages-post *bags-baggage* *instruments-equipment* (*pieces of*) *furniture* *weapons-weaponry* *stationery* (*items*)

Figure 4: Visual stimuli in Experiment 1

4. Testing ambiguity in quantity comparisons. This section reports four experiments that study the hypothesized ambiguity of quantity judgements with CNs and MNs. These experiments introduced participants to situations where quantity judgements are expected to vary depending on whether the comparison is cardinal or non-cardinal. Each situation was accompanied by two comparative sentences as in (6) above. Participants were asked to select one of these sentences, or both of them, as a possible description of the given situation. The experiments involved referents of CNs and MNs whose common perception is as substances (e.g. rock, clay) and discrete objects (bags, apples). The experiments tested Hypotheses (H_1)-(H_2) about the influence of the count/mass distinction on the choice between cardinal and non-cardinal quantity judgements. Specifically:

- Experiment 1 tested exceptional strategies of non-cardinal measurement, comparing OMNs (*baggage*) to CNs (*bags*).
- Experiment 2 tested the effect of a CN/MN environment on exceptional counting strategies with SMNs (*clay*).
- Experiment 3 tested non-cardinal measurement with simple CNs (*apples*) by comparing them to CN phrases that include an overt cardinal expression (*number of apples*).
- Experiment 4 tested counting with simple SMNs (*rock*) by comparing them to SMN phrases that include an overt non-cardinal expression (*volume of rock*).

These experiments are described below in more detail.

4.1. EXPERIMENT 1. The aim of this experiment was to examine if, and to what extent, non-cardinal measurement is tolerated with OMNs more easily than with coreferential CNs, although the referents are preferably perceived as discrete.

4.1.1. MATERIALS AND PROCEDURE. We selected six OMN-CN pairs from (Hampton & Winter 2024):

(7) *packages-post, bags-baggage, instruments-equipment, sofas-furniture, weapons-weaponry, stationery items-stationery*

As in (Hampton & Winter 2024), these nouns were selected using two criteria: (i) minimal referential differences between the OMN and the CN in each pair; (ii) the ease of representing different items, where some items are of a greater volume. For each noun in (7), each participant was presented with a corresponding drawing (Figure 4) and two comparative sentences containing the corresponding OMN (or CN). For example, with respect to the relevant drawing in Figure 4, some participants were presented with the pair of sentences in (8) and others were presented with (9):

(8) a. Anna has more *packages* than Ben. b. Ben has more *packages* than Anna.

(9) a. Anna has more *post* than Ben. b. Ben has more *post* than Anna.

Using sentences and (8-a) and (9-a) to describe the respective situation in Figure 4 reflects *counting*, since Anna has a larger cardinality of packages. Conversely, (8-b) and (9-b) reflect non-cardinal *measurement*. In relation to such pairs, participants were given the following question:

(10) "*Which of the following better describes your reaction to these sentences?*"

Subsequently, participants were asked to choose one of the following two statements:

(11) a. "*I imagine either one of the two sentences might be used to describe the situation*"
b. "*Only one of the sentences can be used felicitously*"

We interpret answer (11-a) as reflecting a judgement about the ambiguity of the given sentence between counting and measurement. Answer (11-b) is interpreted as reflecting lack of ambiguity. Participants who selected option (11-b) received a followup question:

(12) "*Which of the two sentences do you think can be used to describe the image?*"

Accordingly, the answer of each participant was coded as *unambiguous counting*, *unambiguous measurement*, or *ambiguity*.

Using *Prolific*, we recruited 481 speakers of British English (309 female, age M=42.1). Each participant received exactly one of the 12 target questions. In total, between 40-41 responses were collected for each of the 12 items. The experiment started with a 'training' stage. Participants were introduced to the following pairs of sentences concerning a text reporting that Anna's bag weighs 20 kilos, and Ben's bag weighs 10 kilos:

(13) a. Anna's bag is heavier than Ben's bag. b. Ben's bag is lighter than Anna's bag.

(14) a. Anna's bag is heavier than Ben's bag. b. Ben's bag is heavier than Anna's bag.

Answer (11-a) fits (13) and (11-b) fits (14). Participants who answered differently were given an explanation and were asked to correct their answer. This aim of this 'training' was to demonstrate that two sentences with argument reversal as in (8) and (9) may or may not both be true, hence to dissuade participants from adopting automatic strategies when considering such examples.

4.1.2. RESULTS. Table 1 gives the total numbers of unambiguous counting, unambiguous measurement and ambiguity for CNs and OMNs. To test (H_1) we are interested in comparing non-cardinal measurement with OMNs and CNs. For this comparison, we encoded all 'ambiguous' judgements and unambiguous 'measurement' judgements as reflecting the property +MEASURE. In our results, OMNs showed a +MEASURE behavior in 46% of the cases, and CNs in 16% of the cases, yielding a significant effect according to Fisher's Exact Test ($p < 0.00001$). The Odds Ratio was 0.23 (95% Confidence Interval $[0.15, 0.35]$). All OMNs showed a higher frequency of +MEASURE judgements than their corresponding CN, with a significant effect in four of the six pairs. The +MEASURE behavior with OMNs was recorded with levels between 28% and 65%, while with CNs it ranged between 5% and 38%, with only one CN (*stationery items*) showing +MEASURE in more than 20% of the cases. The per item effects are summarized in Figure 5.

4.1.3. DISCUSSION. The results strengthen the claims in (Grimm & Levin 2012, Rothstein 2017) and (Hampton & Winter 2024), according to which OMNs have a privileged access to non-cardinal

	Experiment 1		Experiment 2		Experiment 3		Experiment 4	
	CN	OMN	CN-SMN	SMN-SMN	CN	No. of CN	SMN	Vol. of SMN
counting	202	130	64	27	54	121	27	8
measurement	14	43	55	103	63	29	103	141
ambiguity	25	67	40	31	42	12	31	10
Total	241	240	159	161	159	162	161	159

Table 1: unambiguous counting/measurement and ambiguity in Experiments 1-4

Legend: $[+] : p < 0.05$ $[*] : p < 0.01$ $[**] : p < 0.001$ $[****] : p < 0.00001$

Figure 5: Experiment 1 – 95% confidence intervals of odds ratios, comparing Non-Cardinal Measurement in OMN and CN conditions

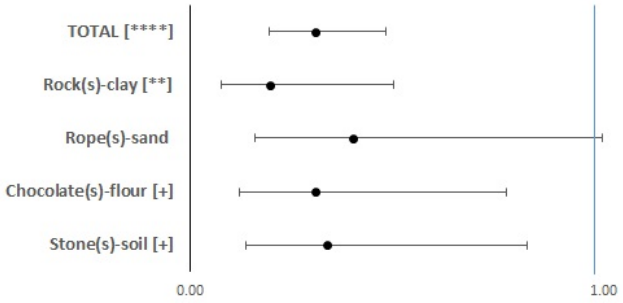

Figure 6: Experiment 2 – 95% confidence intervals of odds ratios, comparing Counting in CN-SMN and SMN-SMN conditions

measurement compared to CNs with the same referents. Notwithstanding, the acceptance of the +MEASURE strategy was not at zero levels with CNs. We may account for this effect as reflecting lack of attention, which led some participants to deviate from the literal meaning of CNs. However, we can also interpret it as reflecting a genuine potential of CNs to trigger non-cardinal measurement. This question will be tested more directly in Experiment 3.

4.2. EXPERIMENT 2. The results of Experiment 1 support the idea that the 'mass' status of OMNs makes their non-cardinal interpretation more accessible than with CNs. OMNs are special among the MNs in that their referents are commonly perceived as discrete objects. However, SMN referents may be packaged into discrete objects like heaps of sand or puddles of water. Do such situations make SMNs more amenable to quantity judgements based on cardinality? The aim of Experiment 2 was to test this question by looking into cardinality judgements with SMNs, and how they are affected by the mass/count distinction in their syntactic environment.

4.2.1. MATERIALS AND PROCEDURE. We selected four nouns, each of them in singular and plural:

(15) *chocolate-chocolates, rope-ropes, rock-rocks, stone-stones*

We selected these nouns because of their flexible SMN/CN use, and, considering the plausibility of comparison with their referents, we selected the following SMNs: *flour, sand, clay* and *soil*,

chocolate-chocolates (vs. flour) *rope-ropes* (vs. sand) *rock-rocks* (vs. clay) *stone-stones* (vs. soil)

Figure 7: Visual stimuli in Experiments 2 and 4

respectively. We used each of the 8 noun tokens in (15) with a pair of sentences like (16) and (17):

(16) a. This image shows more *chocolate* than *flour*. b. This image shows more *flour* than *chocolate*.

(17) a. This image shows more *chocolates* than *flour*. b. This image shows more *flour* than *chocolates*.

These sentences were presented together with the corresponding drawing in Figure 7. Using sentences and (16-a) and (17-a) to describe the situation reflects *counting*, since the image shows a larger cardinality of chocolate candies. Conversely, (16-b) and (17-b) reflect non-cardinal *measurement*. We used the same procedure as in Experiment 1: a question (10) about the pair of sentences, with the two options in (11) and a followup question (12) for participants who selected option (11-b). The experiment started with the same 'training' stage as in Experiment 1.

Using *Prolific*, we recruited 320 speakers of British English (205 female, age M=41.9). Each participant received exactly one of the 8 target questions. In total, between 40-41 responses were collected for each of the 8 items.

4.2.2. RESULTS. Table 1 gives total numbers of unambiguous counting, unambiguous measurement and ambiguity for CN-SMN and SMN-SMN comparatives. To test (H_2), we compared cardinal judgements with SMNs (*flour*) when they appear with SMNs (*chocolate*) and CNs (*chocolates*) as in (16)-(17), encoding all 'ambiguous' judgements and unambiguous 'counting' judgements as '+COUNT'. In our results, comparisons between a CN (*chocolates*) and a SMN (*flour*) showed a +COUNT behavior in 65% of the cases. SMN-SMN pairs (*chocolate-flour*) showed a +COUNT behavior only in 36% of the cases, which is a significant difference from CN-SMN pairs ($p < 0.00001$) according to Fisher's Exact test. The Odds Ratio was 0.30 (95% Confidence Interval $[0.19, 0.47]$). All CN-SMN pairs showed a higher frequency of +COUNT judgements than their SMN-SMN counterparts, with a significant effect in three of the four cases. The +COUNT behavior with CN-SMN pairs was recorded with levels between 53% to 74%, while with SMN-SMN pairs it ranged between 24% and 53%. The per item effects are summarized in Figure 6.

4.2.3. DISCUSSION. The results clearly support hypothesis (H_1) in terms of showing the effect of the CN/MN distinction. Mixed comparisons with CNs (*chocolates*) and SMNs (*flour*) led to significantly more counting compared to their uniform SMN-SMN counterparts (*chocolate-flour*). Effects of the CN/MN distinction on quantity comparisons with 'flexible' nouns also appeared robustly in Barner & Snedeker's experiments. In addition, our results highlight the possibility of 'mismatches' in quantity comparisons: mixed CN-SMN comparatives showed high levels of counting with SMNs (65% of the participants) and non-cardinal measurement with CNs (60%). However, although 'mixed' CN-SMN environments promoted counting with SMNs, also SMN-SMN

comparisons showed a considerable proportion of the participants (36%) who accepted counting: either unambiguously (17%), or in addition to non-cardinal measurement (19%). This shows that when the context makes counting salient (e.g. using heaps of flour as in Figure 7), counting of SMNs is possible even without syntactic pressure. The tolerance of SMNs to counting and of CNs to non-cardinal measurement is studied further in Experiments 3 and 4.

4.3. EXPERIMENT 3. In Experiments 1 and 2, CNs showed non-zero levels of tolerance towards non-cardinal measurement (16% and 60%, resp.). Such effects may be grammatically licensed, but they could also reflect a misinterpretation of the utterance that deviates from its literal meaning. The aim of Experiment 3 was to test to what extent CNs like *apples* are distinguished in this respect from expressions like *number of apples*, where cardinality is explicitly mentioned. If the literal meanings of both expressions trigger counting to the same degree, we expect no difference between them. If the literal meanings of CNs are tolerant towards non-cardinal measurement as in hypothesis (H_2), we expect differences to show up.

4.3.1. MATERIALS AND PROCEDURE. We selected four pairs of CNs:

(18) *apples-almonds, bananas-hazelnuts, cod fillets-peas, potatoes-olives*

These pairs of CNs were selected because of the different sizes of the objects they refer to, and due to their common appearance in recipes (e.g. an apple almond pie). Each of these noun pairs was used with a pair of sentences as in (19) and (20) below:

(19) a. Anna needs more *apples* than *almonds*.
b. Anna needs more *almonds* than *apples*.

(20) a. Anna needs a greater number of *apples* than *almonds*.
b. Anna needs a greater number of *almonds* than *apples*.

These sentences were presented together with the following scenario:

(21) "*For baking an Apple Almond Pie, Anna needs: - 900 grams of apples: about 10 apples and - 70 grams of almonds: about 50 almonds*"

Other pairs of nouns in (18) were studied in a similar way. We used textual representations, as describing recipes using graphical stimuli proved hard. In the situation (21), using sentences (19-a) and (20-a) reflects non-cardinal *measurement*, as Anna needs a larger weight of apples. Conversely, (19-b) and (20-b) reflect *counting*. If measurement is prohibited with CNs, we expect sentences (19-a) and (20-a) to be rejected at similar levels. We used the same procedure as in Experiments 1-2: the same 'training' stage, then a question (10) about the pair of target sentences, with the two options in (11) and a followup question (12) for participants who selected option (11-b).

Using *Prolific*, we recruited 321 speakers of British English (222 female, age M=38.0). Each participant received exactly one of the 8 target questions. In total, between 40-41 responses were collected for each of the 8 items.

4.3.2. RESULTS. Table 1 reports total numbers of unambiguous counting, unambiguous measurement and ambiguity for bare CNs and *number of* CNs. To test (H_2), we are interested in comparing the extent to which speakers apply non-cardinal measurement with CNs in sentences like (19) and (20). For this comparison, we encoded all 'ambiguous' judgements and unambiguous

'measurement' judgements as reflecting the property +MEASURE. In our results bare CNs showed a +MEASURE behavior in 66% of the cases, and *number of* CNs only in 25% of the cases, yielding a significant effect ($p < 0.00001$) according to Fisher's Exact test. The Odds Ratio was 0.18 (95% Confidence Interval $[0.11, 0.29]$). All bare CNs showed a higher frequency of +MEASURE judgements than their *number of* counterparts, with a significant effect in all four pairs. The +MEASURE behavior with bare CNs was recorded with levels between 61% to 77%, while with *number of* CNs it ranged between 20% and 37%. The per item effects are summarized in Figure 8.

Legend: $[+] : p < 0.05$ $[*] : p < 0.01$ $[**] : p < 0.001$ $[***] : p < 0.0001$ $[****] : p < 0.00001$

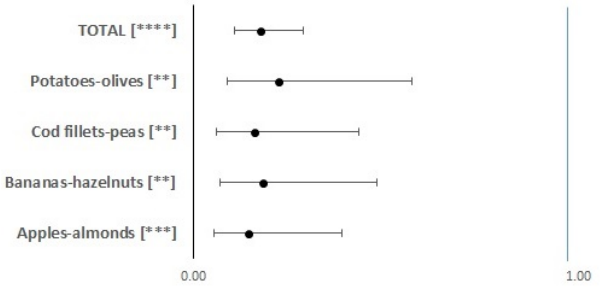

Figure 8: Experiment 3 – 95% confidence intervals of odds ratios, comparing Non-Cardinal Measurement in the CN and *Number of CN* conditions

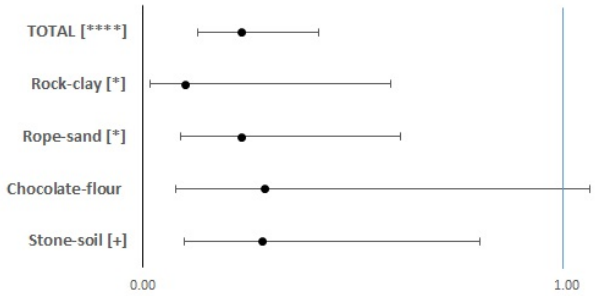

Figure 9: Experiment 4 – 95% confidence intervals of odds ratios, comparing Counting in SMN and *Volume of SMN* conditions

4.3.3. DISCUSSION. Acceptance of non-cardinal measurement was substantially higher with bare CNs compared to *number of* expressions, although the latter also showed considerable tolerance towards measurement (25% of the participants). This supports the assumption that pragmatic pressures for measurement can override literal meaning even with what seems as the clearest example of a cardinal expression. However, this divergence from literal meaning cannot explain why bare CNs showed higher levels of measurement. If the literal meanings of both bare CNs and *number of* CNs categorically rule out non-cardinal measurement, the substantial difference between them remains unexplained. The large difference observed between bare CNs and *number of* CNs strengthens the conclusion that the considerable levels of measurement with CNs in Experiments 1 and 2 are not divergences from literal meanings, but reflect a silent 'grinding' effect, similarly to the use of typical CNs in syntactic 'mass' environments (e.g. *too much rabbit in the chilli*), and in accordance with Rothstein's (2017) analysis of measure expressions with CNs (*2 kilos of nuts*).

4.4. EXPERIMENT 4. Experiment 3 compared non-cardinal measurement with CNs to phrases that explicitly mention cardinality (*number of* CN). Experiment 4 tested the opposite question: to what extent do SMNs show counting effects compared to phrases that explicitly mention non-cardinal measurement (*volume of* SMN)?

4.4.1. MATERIALS AND PROCEDURE. We used SMN-SMN comparatives with the same four pairs of SMNs as in Experiment 2:

(22) *chocolate-flour, rope-sand, rock-clay, stone-soil*

However, each of these noun pairs was now used in pairs of sentences as in (23):

(23) a. This image shows a larger volume of *chocolate* than *flour*.
b. This image shows a larger volume of *flour* than *chocolate*.

These sentences were presented together with the corresponding drawing in Figure 7, where similarly to Experiment 2, sentence (23-a) reflects *counting* and (23-b) reflects non-cardinal *measurement*, The goal was to compare pairs of sentences as in (23) to the four SMN-SMN pairs of Experiment 2, as in (16). We used the same procedure as in Experiments 1-3: the same 'training' stage, then a question (10) about the pair of sentences, with the two options in (11) and a followup question (12) for participants who selected option (11-b).

Using *Prolific*, we recruited 159 speakers of British English. Together with the participants in the SMN-SMN condition of Experiment 2, this resulted in 320 participants (222 female, age M=39.5). Each of these 320 participants received exactly one of the 8 target questions with pairs of sentences as in (16) (Experiment 2) and (23) (Experiment 4). In total, between 39-41 responses were collected for each of the 8 items.

4.4.2. RESULTS. Table 1 reports total numbers of unambiguous counting, unambiguous measurement and ambiguity for bare SMNs and *volume of* SMNs. To test (H_2), we are interested in comparing the extent to which speakers apply counting with SMNs in sentences like (16) and (23). For this comparison, we encoded all 'ambiguous' judgements and unambiguous 'counting' judgements as reflecting the property +COUNT. In our results bare SMNs showed a +COUNT behavior in 36% of the cases, and *volume of* SMNs only in 11% of the cases, yielding a significant effect ($p < 0.00001$) according to Fisher's Exact test. The Odds Ratio was 0.23 (95% Confidence Interval $[0.13, 0.41]$). All SMNs showed a higher frequency of +COUNT judgements than their *volume of* SMN counterpart, with a significant effect in three of the four cases. The +COUNT behavior with bare SMNs was recorded with levels between 24% to 53%, while with *volume of* SMNs it ranged between 3% and 20%. The per item effects are summarized in Figure 9.

4.4.3. DISCUSSION. The acceptance of cardinality comparisons was substantially higher with bare SMNs compared to *volume of* expressions, although the latter also showed some tolerance towards cardinality (11% of the participants). In sentences like (16) (Experiments 2 and 4) and (23) (Experiment 4), the only pressure to apply counting with SMNs was pragmatic (i.e. in the visual stimulus). This is similar to how non-cardinal measurement with CNs was only triggered by pragmatics in Experiment 3. Similarly to our conclusion about CNs in Experiment 3, we interpret this result as indicating that despite the strong contribution of SMNs to the decision on non-cardinal measurement, their literal meaning licenses pragmatic triggering of a 'packaging' operator, similar to the syntactically triggered interpretation of examples like *three beers*.

5. Conclusions. The experiments in this paper show that the mass/count distinction systematically affects the interpretation of comparatives, where mismatches between noun type and quantity judgements are common beyond just object mass nouns. We conclude that all mass nouns encode continuity in their meanings, while all count nouns encode discreteness. Mismatches between meanings and quantity judgements are explained by covert 'grinding' and 'packaging' operators (Rothstein 2017), which can be triggered by syntax or pragmatics. Thus, as our experiments demonstrated, pragmatics may pressure comparatives like *more nuts* or *more clay* to yield an

atypical quantity judgement, despite the lack of mismatch between syntax and lexical mass/count preferences. This fact may have been obscured by eliciting truth-value judgements in neutral contexts, but is revealed when participants are given the opportunity to contribute judgements about ambiguity. We conclude that object mass nouns like *furniture* can be treated as an extreme case of pragmatic mismatches in the mass domain. Aggregate count nouns like *beans, peas* and *lentils* may exemplify the opposite situation: count nouns that pragmatically favor non-cardinal measurement. Further research is needed on cross-linguistic variations with the mass/count encoding of these concepts, and their effects on quantity judgements.

References

Barner, David & Jesse Snedeker. 2005. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition* 97. 41–66. https://doi.org/10.1016/j.cognition.2004.06.009.

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149. https://doi.org/10.1007/s11229-009-9686-6.

Gillon, Brendan S. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15. 597–639. https://doi.org/10.1007/BF00628112.

Grimm, Scott & Beth Levin. 2012. Who has more furniture? Presented at the *Mass/Count in Linguistics, Philosophy and Cognitive Science Conference*. Unpublished ms.

Hampton, James A. & Yoad Winter. 2024. Countability and comparative judgements. https://www.phil.uu.nl/ yoad/papers/HamptonWinterSuB2023.pdf. To appear in *Proceedings of Sinn und Bedeutung 28*.

McCawley, James D. 1975. Lexicography and the count-mass distinction. In *Annual meeting of the Berkeley Linguistics Society*, vol. 1, 314–321. https://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/2335/2105.

Pinto, Manuela & Shalom Zuckerman. 2019. Coloring book: A new method for testing language comprehension. *Behavior research methods* 51. 2609–28. https://doi.org/10.3758/s13428-018-1114-8.

Rothstein, Susan. 2017. *Semantics for counting and measuring*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9780511734830.

Scontras, Gregory, Kathryn Davidson, Amy Rose Deal & Sarah E. Murray. 2017. Who has more? the influence of linguistic form on quantity judgments. *Proceedings of LSA* 2. 41:1–15. https://journals.linguisticsociety.org/proceedings/index.php/PLSA/article/view/4097/3785.

Snyder, Eric. 2021. Counting, measuring, and the fractional cardinalities puzzle. *Linguistics and Philosophy* 44. 513–550. https://doi.org/10.1007/s10988-020-09297-5.

Syrett, Kristen & Julien Musolino. 2015. All together now: disentangling semantics and pragmatics with together in child and adult language. *Language acquisition* 23. 175–197. https://doi.org/10.1080/10489223.2015.1067319.

Winter, Yoad. 2022. Mixed comparatives and the count-to-mass mapping. In *Empirical issues in syntax and semantics 14*, 309–338. http://www.cssp.cnrs.fr/eiss14/eiss14_winter.pdf.