Language Learning and Development

ISSN: 1547-5441 (Print) 1547-3341 (Online) Journal homepage: www.tandfonline.com/journals/hlld20

When and why parts count: linking categorization and counting of partial objects

Athulya Aravind, Karissa Sanchez & Kristen Syrett

LANGUAGE LEARNING AND DEVELOPMENT
https://doi.org/10.1080/15475441.2026.2614449

When and why parts count: linking categorization and counting of partial objects

Athulya Aravind[a,b], Karissa Sanchez[c], and Kristen Syrett[d]

[a]Department of Linguistics & Program in Cognitive Science, Yale University, New Haven, CT, United States; [b]Department of Linguistics & Philosophy, Massachusetts Institute of Technology (MIT), Cambridge, MA, United States; [c]Microsoft, Redmond, WA, United States; [d]Department of Linguistics & Center for Cognitive Science (RuCCS), The State University of New Jersey, New Brunswick, NJ, United States

ABSTRACT

When asked to count using basic-level noun descriptions (e.g. *fork*), children often count both whole objects and "partial objects" (e.g. a broken fork-piece) as one unit, diverging systematically from adults. We test one class of explanations for this "over-counting" behavior: that children apply nouns to objects more permissively than adults, and then count those objects they have labeled with the noun. In a two-task study, we examined how categorization of a partial object under a count noun related to numerical evaluations of sets that included that object, including cases where the numerical expressions involved fractions. Adults categorized and counted partial objects conservatively, often excluding them from the category and counting them, at most, for a fractional share (1/2). Children were more permissive in both tasks: they frequently accepted partial objects as noun category members and counted them on par with wholes (as |1|). Strikingly, children's numerical judgments were often incongruous with their categorization: They sometimes rejected a partial object as a category member, but still counted it as |1|. We argue that while children are indeed more flexible in their noun application, their counting behavior calls for a different account. We propose that children lack the numerical capacities to represent and reason over fractional amounts, and default to a natural number scale even when adults use a finer-grained measurement scale. Children may recognize that the partial fork contributes some non-zero amount to a sum, but the only way they can add it to that sum is as |1|.

Introduction

Some of the earliest words in an English-learner's productive vocabulary are everyday count nouns like *ball* or *cup*. Classic research by Goldin-Meadow et al. (1976) demonstrated that by the age of 2, children understand many common nouns, even when they don't yet produce them. More recent work by Bergelson and Swingley (2012) has shown that infants as young as 6 months of age can associate nouns for body parts or foods with their images, long before these words are ever produced. Given this remarkable early progress, one might expect that by age 4, a learner's treatment of count nouns would have long been solidified. This is an age, after all, when children can speak in full sentences and understand more complex expressions like abstract verbs (Lewis et al., 2017) and logical terms (Chierchia et al., 2001; Gualmini, 2004). But a peculiar type of non-adultlike behavior persists with simple count nouns well into the preschool years. When asked to count or quantify objects that fall under them – words like *ball*, *cup*, or *fork* – children appear to treat whole objects and discrete parts of these objects in the same way, as one instance of ball, cup, or fork.

CONTACT Athulya Aravind athulya.aravind@yale.edu Department of Linguistics, Yale University, PO Box 208366, New Haven, CT 06520-8366, Connecticut

The phenomenon was first documented by Shipley and Shepperson (1990). They presented children with an array of four whole forks and two halves of a fifth broken fork. Children were asked to "count the forks." More than half the time, children reported a total of *six* forks, apparently treating each broken fork-piece as a unit in their counts. Adults, by contrast, either ignored the broken pieces altogether and counted four, or else mentally combined the broken pieces into a whole and counted five. Shipley and Shepperson attributed children's non-adult behavior to a "spatio-temporal" bias in object individuation, which leads children to prioritize discrete physical objects as units for counting. This idea was later expanded on in subsequent work, with Wagner and Carey (2003) arguing that the bias extends to the enumeration of events, and Melgoza et al. (2008) showing that it applies broadly to all forms of quantification and not just counting.

These early theories can be characterized as "conceptual bias" accounts. They begin with the assumption that the correct, adultlike way to count forks is to count only whole, functional objects with the property of being a fork. But children "have a strong bias to use a spatio-temporal individuation strategy when counting objects and . . . will ignore a conflicting linguistic description in favor of this spatio-temporal bias" (Carey & Wagner, p. 163). Spatio-temporal objects are children's "default units" of counting, overriding other individuation criteria, so they count both whole forks and spatio-temporally distinct fork-parts when asked to count forks.

Yet, several findings challenge the view that over-counting arises solely from object-based individuation biases. Giralt and Bloom (2000) showed that children do count entities that are not discrete physical objects, like holes or parts of other objects (e.g., handles, feet). More relevantly, Brooks et al. (2011) found that children were less likely to count parts when those parts, such as wheels separated from a bicycle frame, had names the children knew. These results suggest that over-counting is at least partly tied to children's linguistic knowledge. Building on this idea, more recent proposals (Srinivasan et al., 2013; Syrett & Aravind, 2022) link over-counting specifically to children's understanding and use of nouns.

These "linguistic" accounts diverge from earlier conceptual accounts both in which component of cognition they locate the problem and in their assumptions about developmental continuity. The core idea is that children over-count because they over-*categorize* discrete parts as falling under the extension of the noun. Crucially, adults, too, can use count nouns for partial objects in some contexts. For instance, a shard of pottery from an archeological dig might reasonably be called a *plate*. Thus, applying a count noun to discrete partial objects is not radically discontinuous from adult behavior. The difference between children and adults under these proposals is that children are *overly* flexible, failing to exclude discrete parts even in circumstances where adults limit the noun's use to wholes.

The reasons for children's over-permissiveness with nouns differ across the two accounts. Srinivasan et al. (2013) argue that nouns do not provide the full set of individuation criteria, and do not by themselves exclude partial objects from their extension. To identify individual units, and restrict quantification to whole objects, noun meanings must be *pragmatically* restricted. Specifically, adults restrict unmodified nouns to whole objects alone because they have better and more specific ways of talking about object parts, for example, using measure phrases and other modifiers to describe an object's partial state. In contexts involving partial objects, therefore, more precise alternatives like *piece of fork* outcompete *fork*. But children do not know and/or cannot always spontaneously generate these alternative descriptions for object parts, and thereby allow unmodified nouns to be used in ways that are blocked via pragmatic competition for adults. Supporting this, they found that children benefited from being explicitly given the better alternatives: When asked to choose the better descriptor for two partial shoes, children chose "two pieces of shoe" over "two shoes." The account also straightforwardly accounts for why children don't count detached wheels when asked to count *bikes*. They have ready access to a better alternative: *wheel.*

Syrett and Aravind (2022) offered a different type of pragmatic explanation. They, too, take count noun meanings to be flexible enough to apply to partial objects for both adults and children. Adults are highly attuned to various factors that might restrict the noun's *in-context* application to wholes alone (e.g., object functionality). Syrett and Aravind proposed that children are less able to integrate such contextual information to determine when noun application should be restricted to whole objects of the kind, and when a less restrictive construal is appropriate. In their study, 4-year-olds were less likely to apply a count noun to a partial object when the speaker's goals for the object or the object's intended function were made explicit (e.g., a fork intended for pasta-eating). Thus, contextual support elicited more adultlike behavior, even when there was no change in the available linguistic descriptions. Adults did not appear to be influenced by this contextual information, presumably because they were already using the stricter standard. This account explains children's over-counting in earlier studies as a direct consequence of this over-permissive noun application: Because (absent explicit contextual support) a partial fork *is* a fork just like a whole fork, it also counts for one unit, just like a whole fork.

Despite their differences, both linguistic accounts tie counting behavior to children's command of the count noun. They reduce two non-adult behaviors – allowing partial objects to fall under the extension of a noun in more circumstances than adults, and counting discrete parts on par with wholes – to a single pragmatic source. Thus they both avoid positing any conceptual or lexical-semantic discontinuity between adults and children. Because of this, they also make an important shared prediction: How children count objects should depend, at least partially, on how they categorize such objects. If a partial fork is categorized as *a fork*, it should count as one unit of *fork*; if it isn't categorized as *a fork*, it should be excluded from the count.[1]

The main goal of the present work is to test this shared prediction. We do so in a two-task study that sequentially examines *both* categorization and counting, in order to establish whether the same children who accept a (non-part-denoting) count noun for a partial object are also the ones who count that object for |1|. If, on the other hand, children's categorization and counting were to dissociate, that would imply that over-counting cannot be reduced to noun application (or that the explanation is not solely linguistic), and would motivate a different kind of explanation. We propose one such explanation in the General Discussion.

Beyond the correlational, within-subject design, the study involves two further design features. First, unlike previous work, the counts involve not just whole numbers but also the mixed number "one-and-a-half." The inclusion of this expression is driven by the fact that, in contexts involving partialities, adults often count them according to the fractional shares they represent, rather than simply as 0 or 1 (Bale & Nicolas, 2024; Salmon, 1997).

Second, across the two tasks, we compare behavior on nouns with gradable adjectives. An adjective is gradable if the property it denotes can hold of entities to greater or lesser degrees (Bartsch & Venneman, 1972; Bierwisch, 1989; Cresswell, 1976; Kennedy, 1999; Kennedy & McNally, 2005, a.o.). Within this class of adjectives, there are three main subclasses that differ along their scalar structure and how the standard of comparison is set (Kennedy, 1999, 2007; Kennedy & McNally, 2005; Rotstein & Winter, 2004). *Relative GAs* (REL GAs), like *tall* and *expensive*, have a standard of application that can vary from context to context. What counts as *tall* will differ depending on whether we are discussing 7-year-olds in a school photo or NBA basketball players. But our focus here is on *absolute* GAs, which do not show such contextual variability in standards of application. *Absolute maximum standard GAs* (MAX GAs), like *full* or *straight*, require category members to exhibit the adjectival property to the maximal degree (e.g., a cup must be completely full to be considered *full*, with small exceptions pragmatically licensed). In contrast, *absolute minimum standard GAs* (MIN GAs), like *bumpy* or *striped*, require the mere presence of the property; even possessing a single stripe would qualify an object as *striped*.

We include absolute GAs as a comparison point in our study as they offer informative parallels to nominals with both categorization and counting.[2] These adjectives have clear and distinct criteria for when something can fall under them, which prior research has demonstrated that preschoolers understand (Syrett et al., 2010; Gotowski & Syrett, 2020). Comparing

[1]The predicted correlation is weaker for Srinivasan et al. (2013), who posit only a one-way relationship. We address this difference between the two linguistic accounts in more detail in the General Discussion.

participants' categorization behavior on nominals with absolute GAs thus might clarify the standards that guide their treatment of partial objects. For example, a participant who excludes all partial forks from the category, and categorizes only whole forks as *forks*, might be treating count nouns like MAX GAs, where membership requires the property to its "maximal" degree (in this case, the whole object). Conversely, a participant who allows for fork-pieces to fall under *fork* might be treating count nouns more like MIN GAs, where any degree of the relevant property suffices for category membership[2]. Task 1 of our study thus asks: What degree of "object-ness" must an object possess in order for it to fall under the relevant noun? Is it some minimal degree, or the maximal amount?

Though absolute adjectives support measurement – we can measure out degrees of fullness in volume, just like we can measure the quantity of objects on a numerical scale – their lexical semantics constrain *how* we can do this linguistically. Specifically, measure phrases like *half* cannot be applied to them in the same way as with nouns. While it is possible to count a whole ball and a ball cut in half as "one and a half balls," it is not possible to count a fully straight line and a line bent at the midpoint as "one and a half straight lines." The half-straight line is simply not straight; it is bent. On the most natural interpretation of a phrase like "one and a half straight lines," the scope of the mixed number is the noun *line*: it picks out a set containing one straight line and one half of another straight line. We see something similar with MIN GAs. It would be odd to describe a ball with bumps all over and one with bumps along half of it as "one and a half bumpy balls," because a ball with *any* amount of bumps would simply count as one bumpy ball.[3] The scope is the noun: Differences are also evident in the scope of the fraction when it appear with nouns vs. adjectives: "One and a half balls" is equal to one ball and a half ball. Thus, for participants who command both the lexical-semantics of nouns and adjectives, as well as measurement terms like *half*, the mixed number *one and a half* should pattern differently across the two categories. Task 2, which focuses on numerical judgments, examines how children and adults assess the counts of "one N (one)," "one and a half N (ones)," and "two N (ones)" for a dyad consisting of a whole/maximal and a partial/non-maximal object across categories.

Finally, our within-subject two-task design lets us explore the relationship between categorization and counting by asking: Does a participant's decision to categorize a half-object as an instance of the noun correlate with their cardinality assessment in the second task?

Methods

Transparency and data availability

Stimuli, anonymized data, and analysis code are available on the Open Science Framework (OSF) at: https://osf.io/w92p4/. The child sample of the experiment was pre-registered (OSF pre-registration). Methodological and analytical choices were as indicated there, except when otherwise noted. For the adult sample, we follow identical analysis plans.

[2]Importantly, in utilizing these parallels in our study design, we do not mean to suggest that that gradable adjectives and nouns share identical formal representations (though see Sassoon (2013) for work in this direction). Their representations vary with categorial syntax and language-specific features, and in some languages, gradable adjectives may have different formal-semantic representations from English. Rather, we assume that our conceptual system makes available a shared scalar architecture that can be repurposed across grammatical categories, and which supports measurement of properties, objects, and events, and allows for mapping between them. For example, cross-category parallels have been noted for resultatives (e.g., *hammer the metal flat*) and change-of-state events (e.g., *the sky darkened*), and underlie the notion of incremental themes (e.g., tracking completion of *eat an apple* by the state of the apple).

[3]Of course, pragmatically competent listeners can, in context, construct interpretations of phrases like "one and a half straight lines" that would render such phrases meaningful (for instance, the half-straight line being half the length of the other). Our point is that the lexicalized maximum standard of a GA like *straight* makes at least one critical interpretation unavailable: counting a line that is only partway straight, i.e., a bent line, as one half of a straight line.

Participants

We analyze data from 40 children (mean age = 5;0, range = 4;6–5;7) whose dominant language was English, as well as 40 native-English-speaking adults.[4] Child participants were recruited from local preschools and childcare centers in the Boston area. Adults, primarily university students or young professionals in the Boston area, were recruited via a mailing list for behavioral experiments and compensated with $10 Amazon gift cards. Written consent was obtained from adult participants, as well as parents of child participants.

Stimuli

Stimuli consisted of physical objects designed to represent nominal and adjectival properties to varying degrees. The test items were selected based on imageability and ease of creation, as well as child-friendliness and familiarity of the corresponding expressions involved.

Task 1: Participants viewed sets of seven physical objects exhibiting a property instantiating nominal kinds or adjectival properties in seven stepwise sequential degrees. Properties fell into three categories: nominal (NOM) (*balls, plates, bottles*); absolute minimal GAs (MIN) (*spotted* (flowers), *bumpy* (clay balls), *striped* (rings)), and absolute maximal GAs (MAX) (*straight* (bars), *closed* (boxes), *full* (containers)). At Degree 1, the object did not have the kind/property at all. At Degree 7, it displayed the kind/property to the maximal degree for NOM and MAX, and had a more substantive degree of the MIN property. See Figure 1. We presented a further category (using relative GAs (REL)) exclusively to make sure that participants will show variability across trial types. These REL items (*small* (spheres), *long* (bars), *big* (cubes)) were not part of the analysis and are not pictured.

Task 2: Participants were presented with object dyads from Task 1, consisting of one object displaying the kind/property partially (Degree 4) and one displaying it maximally or to a significant degree (Degree 7). To facilitate item presentation across different orders (see Procedure), instead of using the very same objects as in Task 1, we created exact replicas of Degrees 4 and 7 for each critical category (NOM, MIN, MAX).

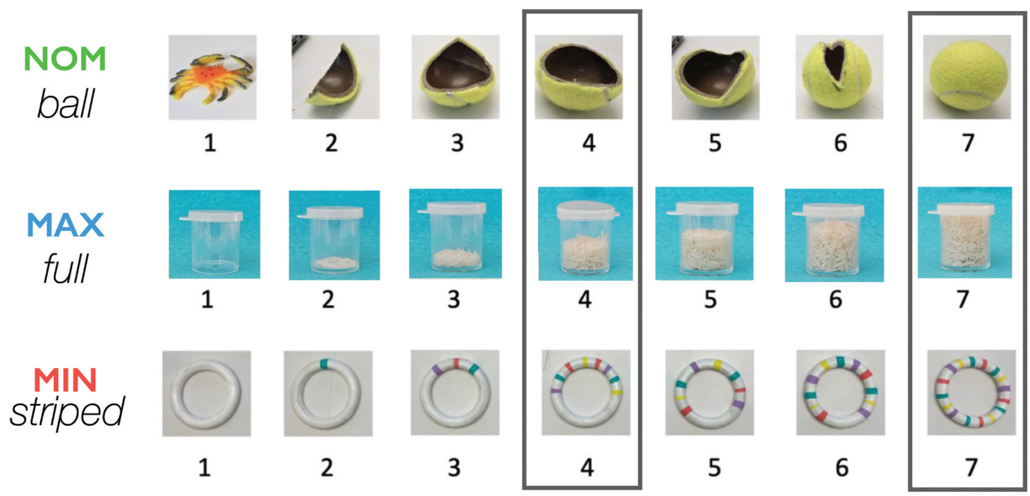

Figure 1. Images of sample items used in Task 1 for each of the three critical categories. Boxed items represent object dyads used in Task 2.

[4]We had originally pre-registered a recruitment age range of 4;6–5;6, but discovered when data collection had completed that one child was in fact 5;7. We decided to use the full 40-child sample. Including or excluding this participant did not alter the significance of any statistical test.

Procedure

Task 1

Participants were shown two felt squares (blue and green) and asked to place "the X (ones)" on the blue square and "the other ones" on the green square (where X was the tested adjective or nominal), thereby creating a partition between category members and non-category members (Gotowski & Syrett, 2020). We only prompted categorization by the relevant positive property X, did not label anything that was not "X," to ensure two partitions without gaps, and to place focus on the target kind/property. The task began with two practice trials (categorizing faces by emotion or animals by kind) to familiarize participants with varying sorting strategies.

Participants then completed partitioning of 12 test items, 3 from each of the critical categories (NOM, MAX, and MIN), along with 3 REL control items. The sets were presented in one of two pseudo-randomized orders.

For each object within a set, categorization decisions were coded as 1 (category member) or 0 (non-member). Our expectations for the adjectival categories were informed by findings from Gotowski and Syrett (2020), who used a structurally similar partitioning task to investigate children's understanding of gradable adjectives. For REL adjectives, we expected the standard to align with the midpoint of the scale, with three to four items assigned to the target square. For MIN adjectives, we expected that the majority of items (six out of seven) would be categorized as "the X ones." For MAX adjectives, we expected only the item displaying the maximal degree of the property to be included, though prior findings suggest that both children and adults may tolerate some deviation from the maximum (see also Syrett et al. (2006)).

Additionally, based on *an individual's* categorization patterns for each MAX and MIN item, we were able to infer the adjectival semantics they assigned to the relevant expression within the task, which we could differ from our a priori classification. Thus, if a participant categorized only maximally (Degree 7) or near-maximally (Degree 6) filled containers as *full*, we coded their category for that adjective as MAX. If instead, they treated any filled container (Degree 2–7) as *full*, we coded their category as MIN. Finally, if the member–non-member split happened somewhere around the halfway point (Degree 4 or 5), the category was coded as REL. This "participant-driven category," rather than some theoretically pre-defined category, serves as the predictor for behavior in Task 2.

Task 2

Immediately following Task 1, participants were introduced to a puppet learning to count. In each trial, the experimenter presented the puppet with an object dyad (one partial, one whole, see Figure 2) comprised of NOM, MAX, and MIN items from Task 1, and asked the puppet how many items were present. The puppet gave a count with one of three verbal measures: "one," "one-and-a-half," or "two" using the sentence frame: "You brought me N NOMs/ADJ ones." Participants were prompted to agree with or (if they disagreed) correct the puppet's counts and provide their own count.

Task 2 began with a practice trial (e.g., ensuring participants corrected the puppet for counting one banana as two). Participants completed nine test items in one of three pseudo-randomized orders. Each participant saw three trials per category, with one trial from each category–measure combination. The pairing of items with measures was counterbalanced across orders: For example, if Participant 1 saw the noun *ball* paired with a puppet's count of "one," Participant 2 saw it paired with "one and a half," and Participant 3 with "two," and so on. Responses were coded as 1 (agreement with the puppet) or 0 (disagreement), with participant-provided counts recorded when applicable.

Results

The results are presented in three parts. We first report findings from each of the two tasks separately. In the final section, we then turn to cross-task comparisons. We conduct separate analyses on child and adult data, as only the child portion of the study was pre-registered. Our primary analyses for the child data adhered to the pre-registered plans, though we also report additional exploratory analyses and post hoc tests where relevant. We follow the same analytic plans for the adult data. All statistical analyses were conducted using logistic mixed-effects modeling with the *glmer* function from the *lme4* package (Bates et al., 2015). Post hoc pairwise comparisons were carried out using the *emmeans* package (Lenth, 2018), employing the Tukey method for multiple comparisons. All data and code are available here: https://osf.io/w92p4/

Task 1

In Task 1, participants classified objects into categories based on the degree to which they displayed a property. With the REL controls, objects began to be reliably included in the category at Degree 5 for both adults and children (88% inclusion for adults, 81% for children), with both groups setting the standard of application based on the comparison class as expected – around the midpoint, splitting the set into groups of three and four objects. These REL adjectives do not figure into our analyses, as we focus on a comparison of MIN, MAX, and NOM, so we now turn to those test items. Figure 2 illustrates the rates at which adults and children categorized objects as members across the three critical categories: MIN, MAX, and NOM. On visual inspection, categorization of NOM objects had a unique pattern, but one that differed for the two populations. It appeared to align more closely with MAX for adults, and more closely with MIN for children.

Our statistical analyses focused on two critical points: Degree 2, the first point where MIN and MAX categories are predicted to diverge (the object should be categorized as a member already for MIN, but not MAX), and Degree 6, the final point at which they are to diverge (the object should still not be a member for MAX). We fit separate mixed-effects logistic regressions on the subset of data corresponding to each of these two Degrees.[5] For each model, the probability of category

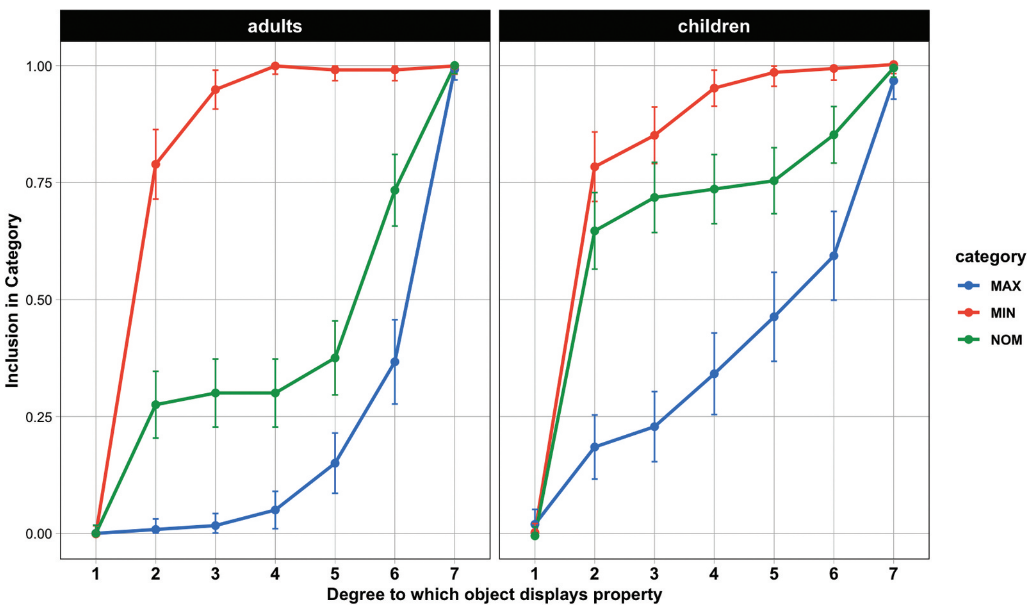

Figure 2. Mean rate of inclusion as a member by category and degree in Task 1; error bars represent 95% CIs.

[5]Our pre-registered analysis plan was to examine differences between Degree 2 and Degree 1, as well as between Degree 6 and Degree 7. However, due to floor and ceiling effects at Degrees 1 and 7 respectively, which made the pre-registered model unimplementable, we decided to focus solely on Degrees 2 and 6.

inclusion (1: included in category; 0: excluded) was predicted using Category (NOM, MAX, MIN) as a treatment-coded fixed effect, with NOM as the reference level, and random intercepts for participants and items.

At Degree 2, adults distinguished NOM from both adjectival categories. Degree 2 objects were less likely to be included in the category under MAX compared to NOM (8.7% for MAX vs. 27.5% for NOM; $\beta = -5.31$, SE = 1.37, z = −3.87, $p < .001$), and more likely to be included under MIN compared to NOM (78.9% for MIN; $\beta = 4.25$, SE = 0.93, z = 4.55, $p < .001$). For children, however, NOM differed only from MAX (18.5% for MAX vs. 64.7% for NOM; $\beta = -2.87$, SE = 0.67, z = −4.28, $p < .001$). Inclusion rates for NOM at Degree 2 was not significantly different from rates for MIN (78.4% for MIN; $\beta = 0.94$, SE = 0.62, z = 1.51, $p = .13$).

At Degree 6, adults again distinguished NOM from both MAX and MIN. MAX adjectives yielded lower inclusion under them at Degree 6 compared to NOM (36.7% for MAX vs. 73.4% for NOM; $\beta =$ −1.9, SE = 0.34, z = −5.64, $p < .001$), while MIN yielded significantly higher rates of inclusion ($\beta = 3.99$, SE = .04, z = 3.84, $p < .001$). Children showed a similar pattern. Degree 6 objects were included as members less often under MAX than under NOM (59.4% for MAX vs. 85.2% for NOM; $\beta = -1.68$, SE = 0.81, z = −2.08, $p = .04$), and more often under MIN than under NOM (99.4% for MIN; $\beta = 3.19$, SE = 1.29, z = 2.48, $p = .01$).

Finally, in a non-pre-registered analysis, we also explored developmental differences in noun and adjective application. We directly compared adults and children by fitting a model to the full dataset, predicting overall inclusion rates (collapsed across Degrees) as a function of Category (treatment-coded as before), Group (adults vs. children; sum-coded), and their interaction. Overall rates of inclusion were comparable across groups (55.4% inclusion rates for children vs. 55.6% inclusion rates for adults), but there were category-based differences. The model found a significant Group by Category interaction at the contrast between NOM and MIN ($\beta = -0.60$, SE = 0.08, z = −7.31, $p < .001$). This was because the difference in category inclusion rates between NOM and MIN was much smaller for children (12.5% difference between categories) than for adults (a 39% difference). The model also revealed an effect of Category at the contrast between NOM and MAX ($\beta = -1.10$, SE = 0.23, z = −4.8, $p < .001$), with MAX yielding lower inclusion rates than NOM in both populations (children: 33.1% for MAX vs. 60.2% for NOM; adults: 29.3% for MAX vs. 49.3% for NOM).

Qualitative data from participants' spontaneous remarks reveal further, more subtle differences between adults and children in their criteria for noun application. When adults included a partial object in the category, they often added hedges noting their "lax" standards. For example, Participant 18 asked, "If they're not complete, do they still count as a ball?" Participant 19 remarked when including partial plates in the category, "It's a plate but not a plate." When excluding an object from a category, adults frequently invoked its function. Participant 32, for instance, rejected a ball at Degree 6, noting, "I was seeing if it bounced and decided no." Participant 36, upon rejecting a Degree 6 bottle, indicated that they felt "a bottle should be closeable."

Children, by contrast, made no mention of function, and even when they allowed the object to be a category member, often remarked on object cohesion, emphasizing the "brokenness" or "crackedness." Interestingly, children often used the term "half" to describe partial objects, even when the item in question was not, strictly speaking, half of a whole, hinting that they took "half" to mean something like a "portion" or "part" of X. For example, one child referred to a ball at Degree 2 as "half," while another described a ball at Degree 6, with just a small piece missing, as "half." We return to this observation in our discussion of Task 2.

Task 2

In Task 2, participants evaluated whether the puppet's count of a partial–whole object dyad was accurate for one of three possible counts: "one," "one-and-a-half," or "two." We use the participant-driven adjectival categories from Task 1 to inform their responses in Task 2. Participants' treatment of the adjectives conformed to theoretical expectations 85.7% of the time for adults and 64.6% of the time for children. In both cases, mismatches were due to participants sometimes treating the *absolute* MIN or MAX adjectives more like *relative* adjectives, setting a contextually based standard rather than the expected maximal or minimal standard. Consistent with prior work (Syrett et al., 2010, Gotowski & Syrett, 2020), for both groups, this behavior was most common with the adjective *full* (57.5% agreement for adults, and only 17.5% for children), which often got re-interpreted by children as *filled*. Trials where the adjective was not categorized according to theoretical endpoints (potentially categorized as REL) were excluded from further analyses. This approach resulted in some data-loss (33 observations from the adult data-set, and 65 observations from the child data-set), but was necessary, as we did not have hypotheses about how children (or adults) should count our adjectival dyads if they were relying on contextually varying standards.

Figure 3 shows the rates at which participants accepted the puppet's counts across categories and measures. For adjectival categories, both adults and children responded largely as expected. Both populations agreed with measures of "two" for MIN, but rejected "one." By contrast, both populations tended to agree when the puppet used "one" for MAX adjectives and disagree with "two." This pattern makes sense: A half-full item is *not* a full one, and a half-striped item *is* a striped one and should count for "1." *Neither* group reliably agreed with the measure "one-and-a-half" for either adjective. However, we will later suggest that the low agreement rates for the mixed number stem may have stemmed from different sources for children and adults.

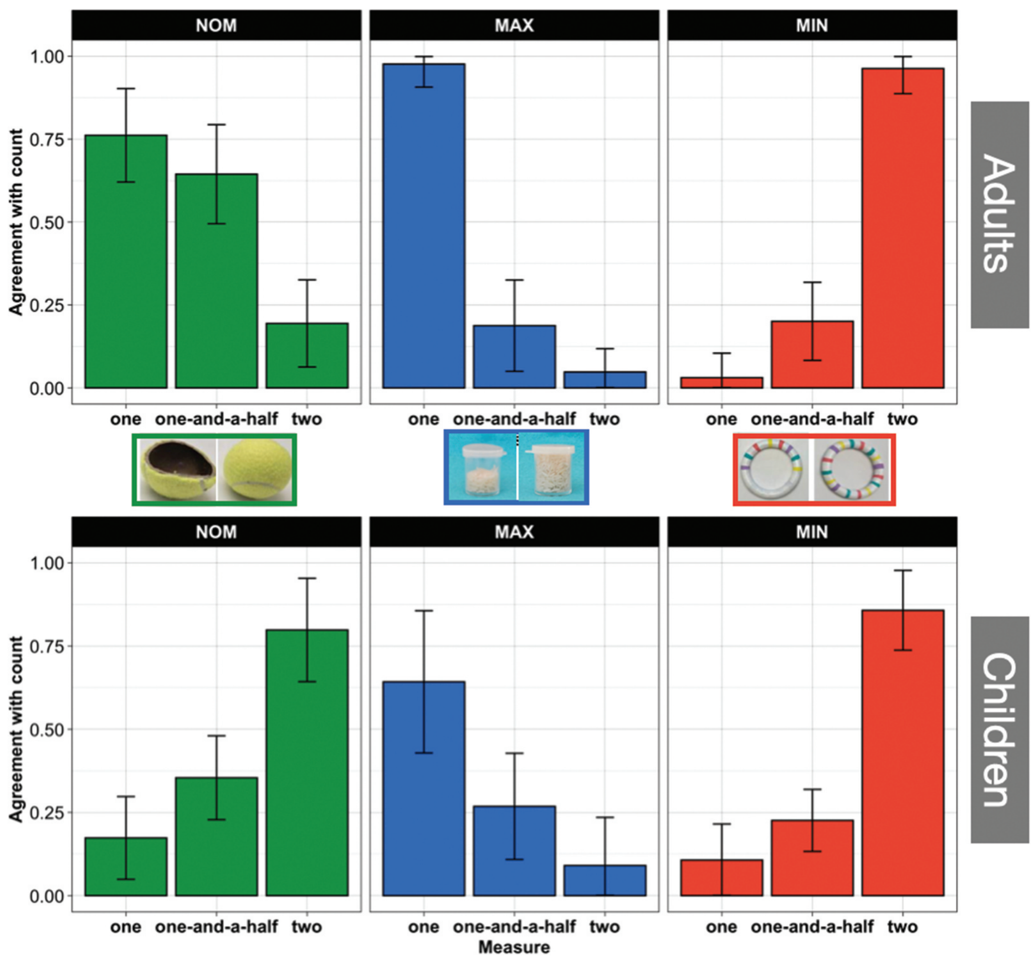

Figure 3. Rate of agreement with puppet's count by category and measure in Task 2; error bars represent 95% CIs.

To analyze these patterns statistically, we fit separate mixed-effects logistic models for each population. We predicted agreement with the puppet from two predictors (Category and Measure) and their interaction as fixed factors, using random intercepts for each participant. Category was treatment coded with NOM as reference level. We helmert coded Measure. This allowed us to explore two critical contrasts. First, the one vs. one-and-a-half contrast shows whether mixed number is accepted when the set includes a partially suitable item. Second, the two vs. less-than-two ("one" and "one-and-a-half") contrast probes whether participants "over-counted" the partial item for a full item.

For adults, this model revealed a significant difference between two and less-than-two at NOM, with "two" being less likely to yield agreement ($\beta = -0.80$, SE = 0.17, z = −4.76, $p < .001$). Additionally, we found three significant interactions: (i) Category at the NOM-MAX contrast and Measure at the one vs. one-and-a-half contrast ($\beta = -2.24$, SE = 0.63, z = −3.56, $p < .001$). This illustrates that for NOM "one" (76.2%) and "one-and-a-half" (64.4%) were both accepted at comparable rates, while for MAX there was a large difference between the two, with "one" (97.6%) being much more frequently accepted than "one-and-a-half" (18.8%). (ii) The second interaction was Category at the NOM-MIN contrast and Measure at the two vs. less-than-two contrast ($\beta = 2.87$, SE = 0.44, z = 6.51, $p < .001$). This interaction reflects that "two" yielded higher agreement rates with MIN (96.3% for two vs. 11.6% for less-than-two), while measures less than two yielded higher agreement rates with NOM (19.4% for two vs. 70.3% for less-than-two). (iii) Finally, Category at the NOM-MIN contrast significantly interacted with Measure at the one vs. one-and-a-half contrast, due to the fact that the difference between the two categories was less pronounced at "one-and-a-half" (a 44.3% difference) compared to "one" (a 73.1% difference; $\beta = 1.38$, SE = 0.61, z = 2.26, $p = .02$). Post hoc pairwise comparisons carried out using the *emmeans* package showed that at "one" and "two," NOM and MAX patterned similarly, and unlike MIN. At "one-and-a-half," however, NOM diverged from both adjectival categories. See Table 1 (left) for the full set of contrasts.

The child model revealed a significant effect of Measure on NOM at the two vs. less-than-two contrast, but in the opposite direction of adults: Counts of "two" were more likely to yield agreement (85.7% agreement at "two" compared to 26.4% at less-than-two; $\beta = 0.89$, SE = 0.17, z = 5.13, $p < .001$). Two additional interactions were observed: (i) Category at the NOM-MAX contrast and Measure at the one vs. one-and-a-half contrast: The difference between NOM and MAX was greater at "one" compared to "one-and-a-half" and in opposite directions (46.9% *higher* agreement for MAX at "one" vs. 8.6% *lower* agreement at "one-and-a-half"; $\beta = -1.53$, SE = 0.48, z = −3.20, $p = .001$). (ii) Category at the NOM-MAX contrast and Measure at the two vs. less-than-two contrast: NOM was more likely to yield agreement at "two" (79.8% agreement with NOM vs. just 9% with MAX), whereas MAX was more often accepted at measures less than two (45.5% agreement with MAX vs. 26.35% with NOM; $\beta = -1.82$, SE = 0.41, z = −4.41, $p < .001$). Post hoc tests (see Table 1; right) showed that children distinguished NOM from MAX at both "one" and "two" and differentiated the adjectival categories from one another at these measures. Critically, children did not distinguish across categories for "one-and-a-half" – a pattern markedly different from what was found with adults.

Table 1. Post hoc pairwise tests of cross-category differences at each measure.

		Adults			Children		
		Estimate	*p*	Sig	Estimate	*p*	Sig
one	NOM – MAX	−2.29	0.09	n.s	−2.43	<.001	***
	NOM – MIN	4.88	<.001	***	0.82	0.51	n.s
	MAX – MIN	7.17	<.001	***	3.26	<.001	***
one-and-a-half	NOM – MAX	2.18	<.001	***	0.63	0.64	n.s
	NOM – MIN	2.13	<.001	***	0.67	0.39	n.s
	MAX – MIN	−0.05	0.996	n.s	0.04	0.998	n.s
two	NOM – MAX	2.10	0.13	n.s	4.56	<.001	***
	NOM – MIN	−5.10	<.001	***	−0.55	0.67	n.s
	MAX – MIN	−7.20	<.001	***	−5.1	<.001	***

To further explore this difference, we fit a final model that compared across children's and adults' behavior on "one-and-a-half." The model predicted agreement from group (adults vs. children; sum-coded) and Category (helmert coded to contrast (i) NOM vs. adjectives, and (ii) MAX vs. MIN) and their interaction as fixed factors and random by-participant intercepts. The intercept coefficient for this model was significant and negative ($\beta = -2.07$, SE = 0.71, z = −2.91, $p = .003$), suggesting that the overall agreement rates for "one-and-a-half" was below chance. There was a significant positive effect Category at the NOM vs. adjectival contrast ($\beta = 0.97$, SE = 0.26, z = 3.66, $p < .001$), telling us that participants were more likely to agree with the measure in NOM compared to the adjectival categories. However, this category effect interacted with group ($\beta = -0.48$, SE = 0.20, z = −2.38, $p = .02$), with children showing much weaker distinctions between NOM and the adjectives (12.5% higher agreement in NOM than the adjectives) compared to adults (45.9% higher agreement in NOM).

Overall, there is no evidence that children have an adultlike understanding of "one-and-a-half." While they do quantitatively differentiate between NOM and adjectives in relation to this measure, this behavior may simply reflect their certainty with counting objects that possess the adjectival properties, and their understanding that *half* means something other than whole (i.e., they know the measurement cannot be *one-and-a-half* striped ones because there are clearly *two* striped whole objects). Indeed, children's spontaneous justifications for disagreements with the puppet support this interpretation. When shown a dyad of partial–whole balls, prompted with "one and a half," Participant 5 corrected the puppet's counts to "two" and went on to remark that "one and a half is *too many*." Participant 37 noted that "one and a half would be *half of one*." The apparent lack of understanding of the mixed number, furthermore, may be connected to their spontaneous, non-adultlike productions of "half" in Task 1. It seems plausible that children only have an interpretation of "half" as a marker of incompleteness or degradation (and thus usable even when the object in question is clearly less than half in quantity), and crucially, not as a measurement expression that is compatible with number meanings.

Comparing across tasks

Finally, we explored the relationship between categorization in Task 1 and numerical judgments in Task 2. Specifically, we analyze how participants' treatment of the partial object at Degree 4 in their Task 1 categorization correlated with their responses to the measure provided by the puppet in Task 2. As a reminder, in Task 1, adults predominantly *excluded* the Degree 4 object from the noun category (30% inclusion rate), whereas children tended to *include* it in the category (74.1% inclusion rate).

Based on participants' responses in Task 2, we created a variable called "Participant Count." This variable aimed to establish what the participant believed to be the cardinality of the set. If a participant agreed with the measure that the puppet gave, that measure was treated as the "Participant Count" for that trial. Thus, if a participant agreed with a puppet's count of "two," the "Participant Count" was recorded as "two." If they disagreed with the measure given, they were always asked to correct the puppet. The value they gave in the correction was recorded as the "Participant Count" in this case. Thus, if a participant disagreed with a puppet's count of "two" and corrected it to "one," we recorded it as "one." Note that a single value (e.g., "two") can appear more often than the number of trials involving that measure alone. This occurs because both agreeing with "two" in a "two" trial and correcting the puppet to "two" in a "one-and-a-half" trial would each be coded as a Participant Count of "two." Figure 4 shows the proportions of each such Participant Count, based on the treatment of the partial object in Task 1 (included in category vs. not).

We begin with a description of the trends, before turning to the statistical analysis. Adults who included the partial object as a category member in Task 1 gave more counts of "one-and-a-half" (at 58.1%) than "two" (27.9%). Still, this group was more likely to give counts of "two" more than those who excluded the partial object from the category, who gave counts of "two" 1.3% of the time. The excluders tended to favor a count of "one" for the partial–whole dyad, agreeing with or offering that count 77.6% of the time, though they also sometimes accepted "one and a half" (19.7%). Children

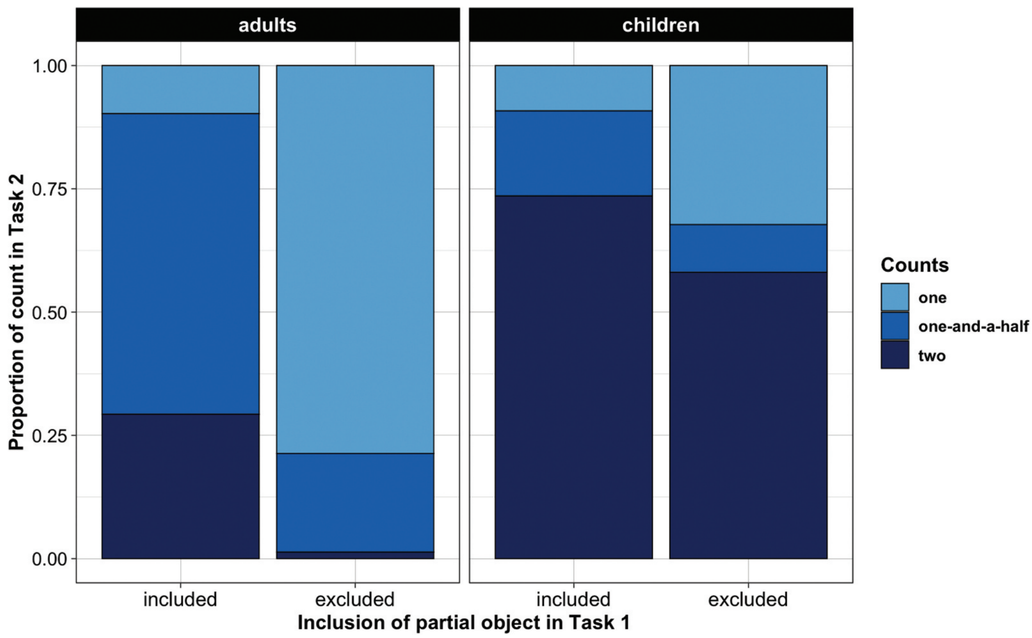

Figure 4. Proportion of each type of count in Task 2 based on whether the Degree 4 partial object was included or excluded from the nominal category in Task 1.

seemed to exhibit a wholly different pattern. Those who *included* the partial object in the category gave the count "two" most often (72.7%), as one might have anticipated. Interestingly, even children who *excluded* the partial object from the category tended to agree with, or offer, counts of "two" (58.1%), and critically, more often than they gave counts of "one" (32.3%).

We fit two models exploring these trends. Our primary statistical question was how inclusion/exclusion of the partial object from the category in Task 1 affected over-counting (i.e., agreeing with counts of "two"). We therefore fit a model predicting the probability of endorsing "two" based on partial object inclusion in Task 1 (coded as 1 for inclusion and 0 for exclusion), group (adults vs. children), and their interaction, with random intercepts for participants. Both factors were sum-coded. The model revealed a main effect of group ($\beta = -2.09$, SE = 0.41, z = 5.12, $p < .001$), showing that adults were less likely overall to endorse "two" (14.6%) compared to children (65.4%). Importantly, there was a significant interaction between group and partial object inclusion ($\beta = 0.72$, SE = 0.33, z = 2.18, $p = .03$), driven by the fact that including the partial object as a category member in Task 1 increased the likelihood of endorsing "two," but primarily for adults. By contrast, for children, the effect of inclusion was much weaker; children tended to endorse "two" regardless of whether the partial object was included in the category or not. Thus, to the extent that adults count a partial object on par with a whole object, they do so in a way that is tightly linked to their inclusion of the partial object as a category member in Task 1. In contrast, categorization decisions in Task 1 seemed to be less strongly correlated with children's counting behavior: They tended to over-count irrespective of whether they initially categorized the partial object as falling under the noun.

Second, we asked how inclusion/exclusion of the partial object in Task 1 affected rates of giving a mixed number count in Task 2. Our second model therefore predicted the probability of endorsing "one-and-a-half" based on partial object inclusion, group, and their interaction, with random intercepts for participants. This model, too, revealed differences both of age group and Task 1 behavior. Adults were significantly more likely to endorse "one-and-a-half" overall (39.9%) than children (8.6%) ($\beta = 0.72$, SE = 0.22, z = 3.27, $p = .001$). This is consistent with the possibility that many of our child participants simply did not know what this expression meant. Inclusion of the partial object in the category had a significantly positive effect on endorsing "one-and-a-half" ($\beta = 0.61$, SE = 0.21, $z = 2.90$, $p = .004$), indicating that categorizing it under the noun does not force it to be counted strictly as |1|.[6]

In sum, for both groups, there seems to be at least partial dissociation of categorization and numerical judgments, but in different ways. Adults sometimes allowed the partial object to be a category member in Task 1, but recognized its partial status in Task 2 and thus accepted a whole number count of "two" only around a quarter of the time. However, regardless of how children categorized the partial object in Task 1, they gravitated toward including it in the count of "two" in Task 2. Thus, partial or whole, the object counts. This discrepancy between categorization and counts is a challenge for the current linguistic accounts, as we elaborate in the Discussion.

Discussion

In this paper, we explored how children's and adults' categorization of discrete parts relate to their count judgments in a correlational within-subject design. Results from Task 1 showed that adults tended to categorize nominals similarly to MAX adjectives, albeit with slightly more flexibility: Nearly whole objects were also included under the noun. In contrast, children categorized nominals more like MIN adjectives but with added stringency: As long as an object possessed some degree of noun-ness, it could fall under the extension of the noun.

Results from Task 2 revealed further differences between children and adults. Adults counted partial objects cautiously, often excluding them entirely and only accepting "one" or counting them for fractional shares. Consistent with this, they frequently accepted either "one" or "one-and-ahalf," but not "two," as an appropriate measure for the whole–partial dyad. Children, in contrast, did not differentiate between partial and whole objects in their counts, echoing the "over-counting" pattern described in previous research. Thus, they primarily accepted counts of "two" for the nominal dyad. Furthermore, they were less likely than adults to endorse counts of "one and a half."

Finally, when we examined the relationship between partial-object categorization in Task 1 and counting of whole–partial dyads in Task 2, we found that count judgments were partially independent of categorization decisions for both groups, but in different ways. Adults were hesitant to count a partial object as |1|, even if they had categorized it as an instance of that noun. Conversely, children often counted the partial object for |1| regardless of whether they had previously included or excluded it from the noun category: For the purposes of counting, once again, partial objects count as wholes. This pattern for both groups demonstrates that categorization does not straightforwardly drive counting judgments.

Implications for previous accounts

Summing across tasks, we find clear differences between adults and children in their categorization, counting, and how they construe the relationship between the two. First, in the categorization task, children differed from adults in degree rather than kind: They were more likely to include partial objects under the noun category, but adults did so occasionally as well. Both linguistic accounts that motivated the present work (Srinivasan et al., 2013; Syrett & Aravind, 2022) assume adultlike lexical semantics in children, but can still accommodate this pattern. For Srinivasan et al. (2013), the child–adult difference arises because children have difficulty accessing better descriptions for parts, and this may have led to their over-permissive categorization. On Syrett and Aravind's (2022) account, children are less sensitive to the contextual cues that adults use to

[6]As noted, the data we analyzed (Participant Count) combines two components: participants' agreement with the puppet's count and the count they provided when they disagreed. Adults' preference for "one-and-a-half" is evident even when considering agreement rates alone: When participants included the partial object, they accepted the puppet's statement with "one-and-a-half" 77% of the time, and even when they excluded it, they accepted it 57% of the time. The lower rates seen in the Participant Count measure reflect the fact that adults spontaneously produce "one-and-a-half" as a correction less frequently.

constrain noun application to wholes; this, too, could have led to them being more inclusive of partial objects in the category.

In the numerical judgment task, there are two main ways children differed from adults. First, children preferred counts of "two" for sets comprising one whole object and one partial object.

This is, of course, the well-documented pattern of over-counting, beginning with Shipley and Shepperson (1990). A second, novel finding is that children struggle with mixed numeral expressions such as "one and a half." This finding is not predicted on either of the linguistic accounts per se. In fact, Srinivasan et al. (2013) report findings that seem to suggest that children both understand "half," and prefer descriptions involving "half" over unmodified nouns when referring to partial objects.

One possible explanation for the divergent findings is the polysemy of the word *half*. In addition to its "quantificational" use as a precise unit of measurement (1⁄2), it also has a loose "modifier" meaning of attributing the property of partiality to something (e.g., *half-moon*, where it is an adjective; *half-eaten* where it is an adverbial). This modifier use is both syntactically and semantically distinct from the use of (*one/a) half* as a precise unit of measurement, and in fact, the two can co-occur: "Give me exactly half of your half-apple." Children may command the adjectival or adverbial use of "half," but still not know its use as a unit of measurement. In our study, too, children spontaneously produced "half" and other modifiers (e.g., *piece*, *broken*) to describe partial objects, suggesting they had mapped some meaning to the expression. But this use case is not evidence of them understanding the quantificational meaning of "half," which is the meaning that is needed when the term appears in mixed numeral expressions like "one and a half." An additional point to note is that since children *do* seem to have the adjectival meaning of "half," "half" was contextually salient in our task, and a noun modified with "half" is a more precise description of the partial object, the behavior we observe is unlikely to arise if access to better alternatives were the only factor at play.

Finally, our critical finding was that categorization does not drive counting behavior. Adults who categorize leniently (as most children) do not consistently over-count, and children who categorize stringently (as most adults) still over-count. These results are difficult to reconcile with existing linguistic accounts. Noun over-application is central to Syrett and Aravind's explanation of over-counting, so neither the child nor the adult pattern is straightforwardly explained. Srinivasan et al. fare better, as they predict only a one-way association. On their view, noun application underdetermines counting judgments. In particular, they propose that pragmatic processes can restrict which elements should count from the set of entities that the noun applies to. Thus, the adult pattern, where an object that was initially included in the category fails to be counted as |1|, is compatible only with their account. What remains harder to explain on this view is the child pattern: Children appear to be counting something as |1| even when they previously found the noun not to apply to it. Pragmatic restriction as a process can straightforwardly output a smaller set than the set of all entities that the noun applies to, but it is unclear how it could output a larger set.

While it might be possible to extend the linguistic theories in ways to capture the observed dissociations, here we take a different path and propose a novel account, one that places children's difficulties not in finding the right conceptual or linguistic description for partial objects, but in understanding how such partial objects should be measured out as fractional quantities.

A new account: non-adult measurement

When it comes to individuation (or having dinner for that matter), a half-fork might not be a fork at all. But when it comes to *measurement*, partial objects *do* count, and represent fractional shares: A half-fork might not be |1| fork, but it has to be more than 0. Adults have access to a conceptual repertoire that includes rational numbers, which allows them to precisely represent, compare, and perform arithmetic over non-integer quantities. When it comes to quantificational contexts involving partiality, they often recruit these dense measurement scales, enabling them to count things like "two and a half oranges" (Bale & Nicolas, 2024; Salmon, 1997). Our proposal is that children over-count because they lack these conceptual/numerical abilities. In particular, they may recognize that a broken half of a fork is a fork-piece, distinguish it from a whole fork, and even understand that it can form a whole with another piece, but still cannot conceptualize precisely *how much* it is. So, when counting, they rely on an integer scale, even when adults opt for a finer-grained scale.[7] They may recognize that half a fork puts the *amount* of fork over zero, but without the tools to measure out ½ or do calculations like 1 + ½ = 1.5, their only available strategy for adding it into a sum of forks is as |1|.

This also is a "conceptual" account, but one that identifies the issue in children's being unable to measure out exact quantities between 0 and 1. Thus, rather than being biased to think that a broken fork-piece is an instance of FORK, on this account, children would think that any quantity μ of FORK, where $0 < \mu^{3} 1$ counts for |1| FORK. This would be a novel explanation for *why* children's "default unit" for counting seem to be spatio-temporal objects: Every spatio-temporal object with some FORK property happens to fit the relevant quantity measurement.

There are clear ways in which this "measurement" account relates to and differs from "linguistic" accounts as well. If children lack the representational capacity for fractions, this should lead to difficulties comprehending word meanings that map onto fractional quantities, such as "half" or "third." Expressions like "half a fork" or "a third of a shoe" not only provide a precise quantity of the relevant objects, they also serve as better *descriptions* of the partial fork or shoe than unmodified nouns like "fork" or "shoe," just as Srinivasan et al. (2013) claim. For Srinivasan et al., difficulties accessing these better descriptions is itself the cause of over-counting. But on the measurement account, what children need is not better descriptions for the partial objects, but better *ways of measuring* their contribution to a sum. The measurement account is thus compatible with the child having access to better descriptions for partial objects, like "broken," "piece," or "part," or even words like "half" used adjectivally, but still show non-adult counting behavior because they lack the precise numerical concepts required for partiality measures.

On this account, then, children's over-counting behavior is a fallback solution to a particular conundrum: They want to count the partial object somehow, but they lack access to the right type of numerical representation for counting fractional shares. But the account is compatible with other ways of getting out of the conundrum besides over-counting. For instance, they could decide that the partial object could be counted a *different* noun. This is one way to think about studies where children do not count discrete parts with known labels, for instance, when they do not count bikewheels when counting bikes. Categorizing the partial object differently (e.g., as falling under *wheel* rather than under *bike*) is another solution to the challenge of having to add up a bike-wheel's share to a quantity of bikes.

Beyond the ability to explain the over-counting behavior (which other accounts can also do), the non-adult measurement account offers a straightforward explanation for children's difficulties with measure phrases like "one-and-a-half." Without the capacity for representing the quantity that the mixed number represents, they cannot learn expressions that refer to such quantities. One piece of evidence in support of this explanation is the way children interpreted the word "half" in the study. Both the quantitative patterns observed with the mixed number and qualitative observations suggest that children's treatment of "half" is more limited than that of adults, and crucially, does not reflect its use as a precise measurement term.

Perhaps the primary advantage this account has over the linguistic accounts, at least when it comes to explaining children's over-counting, is that it does not predict a correlation between noun application and numerical judgment. It is compatible with children having the same underlying nominal semantics, and even fully mature pragmatics, but still counting arbitrary parts, because these actions may rely on different – and differently developing – systems: The former drawing on numerical concepts involving fractions, and the latter drawing on category membership and count noun

[7]Critically, we are not claiming that adults *always* access this fine-grained scale (in fact, they do not always do so in our study). Rather, adults have access to both scales, and therefore a wider range of options, whereas children have only one.

semantics. Whether an object might be reasonably referred to as "a fork" is at least partially independent of the question of how to measure and do arithmetic over fractional fork-shares.

The dissociation of categorization and counting in adults is also succinctly explained on this account. In our study, adults generally accepted the mixed number "one-and-a-half" as an appropriate cardinality measurement for the partial–whole dyad, and furthermore, preferred this finer-grained way of counting the partial object even when they were ready to treat it as a category member. The reliance on fractional reasoning can also help explain adults' counting behavior in previous studies. For example, Srinivasan et al. (2013) found that 88% of the time, adults in their counting task counted a sock cut into two pieces as one sock. This suggests that adults recognize each partial sock as representing a fractional share of 1/2, and when these parts are combined, they total exactly, and no more than, |1|. But it is precisely this type of calculation that we suggest children at this age cannot do. They may very well recognize that the sock pieces can fit together to form a whole sock, but they lack the numerical representations of the part–whole relation, i.e., fractions, needed to determine how much each piece contributes to a sum.

Prior evidence that might support the measurement account

To our knowledge, this hypothesis of discontinuity in measurement concepts has not been explicitly put forth in the literature on partial object counting, but there is suggestive evidence from the number cognition literature that supports the idea (e.g., Gelman, 1991; Sophian, 1997; Hartnett & Gelman, 1998; Mix et al., 1999; Smith et al., 2005; Boyer et al., 2008). Difficulties with fractions are well-documented in school-aged children (Behr et al., 1984), who often make errors that reveal misunderstandings about basic fraction properties, such as the relationship between the numerator and denominator or the relative size of fractional parts. While part of this difficulty may stem from understanding the notation (e.g., why fractions are represented with two numbers), researchers have suggested that it may also reflect a late-developing conceptual representation of rational numbers, essential for thinking about numerical concepts like ¾ or 0.9 (Carey, 2009; Gelman, 1991).

Fraction calculation abilities lag behind the ability to do arithmetic with whole numbers (Gelman, 1991; Mix et al., 1999), with even the most optimistic estimates placing fraction competence at around 6 years of age (Goswami, 1989; Spinillo & Bryant, 1999; Sophian, 2000; Boyer et al., 2008). In a study conducted by Gelman and colleagues (Gelman et al., 1989), which involved reasoning using a special number line (one without numbers, or "cardinal numergraphs"), 5- to 7-year-olds behaved as though they did not understand that numbers can exist between integers. Children were asked to place circle representations of fractions on a special number line schematized in Figure 5, where whole numbers were marked with sets of circles. Many children were categorized as "whole number placers," who placed fractions such as ¼ or ½ at the "one" position, and cards with 1½ at the "two" position. Both the task and children's behavior in it bears a striking resemblance to tasks that probe children's treatment of partial objects. Both involve determining the amount of objects (e.g., forks, circles) that are in a set

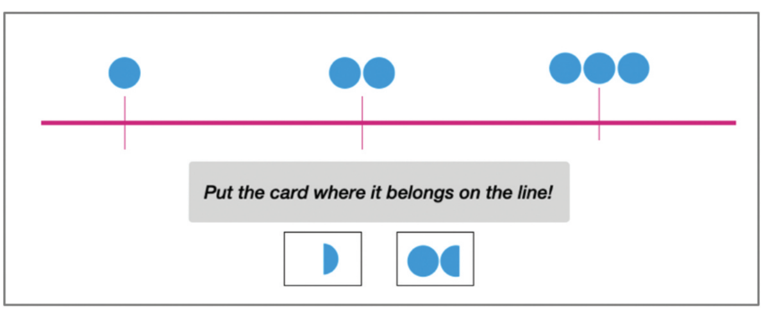

Figure 5. Schema of number line task in Gelman et al. (1989).

comprised of wholes and parts. And in both cases, children do not treat fractional shares in an adultlike manner, failing to appreciate their partial status.

Limitations and future work

We finish on two limitations of our present exploration that open new directions for future research. The first one is methodological. As most children were lenient in categorization, we had a limited number of observations from children who were stringent categorizers, even though this group was important in finding the disconnect between categorization and partial object counting.

However, prior work offers a possible solution to be explored in future work. Syrett and Aravind (2022) demonstrated that when contextual goals for the noun in question were highlighted, children were less inclined to categorize partial objects as falling under that noun. With a similar contextual manipulation future work could create a robust sample of stringent categorization from this population. Such context could provide an economical way to further test the relationship between categorization and counting.

The present work is also limited in that the evidence for the non-adult measurement account is somewhat indirect. While it provides a plausible explanation for the disconnect between categorization and counting, children's knowledge of fractions was not directly tested, except in the interpretation of a mixed number term ("one and a half"). Future work could probe more direct predictions of the non-adult measurement account. One such prediction is that children's measurement difficulties extend beyond nouns and objects, affecting the measurement of fractional shares of continuous quantities (e.g., half a teaspoon of sugar) or non-atomic quantities (e.g., half of a pair of shoes). An equally interesting extension would be to cases where representational units are bounded not by physical properties, but by temporal ones, for instance events (see, e.g., Ji & Papafragou, 2020; Mathis & Papafragou, 2022; Wagner & Carey, 2003). Finally, a clear prediction across domains is that access to fractional shares (independently measured or even elicited by training) should lead to more adultlike counting.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the National Science Foundation [BCS-2016963/2016895].

References

Bale, A., & Nicolas, D. (2024). Counting individuals and their halves. *Linguistics and Philosophy*, *47*(5), 867–914. https://doi.org/10.1007/s10988-024-09418-4

Bartsch, R., & Vennemann, T. (1972). The grammar of relative adjectives and comparison. *Linguistische Berichte*, *20*, 19–32.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., Green, P., & Bolker, M. B. (2015). Package 'lme4'. *Convergence*, *12*(1), 2.

Behr, M. J., Wachsmuth, I., Post, T. R., & Lesh, R. (1984). Order and equivalence of rational numbers: A clinical teaching experiment. *Journal for Research in Mathematics Education*, *15*(5), 323–341.

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bierwisch, M. (1989). The semantics of gradation. *Dimensional Adjectives*, *71*(261), 35.

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology*, *44*(5), 1478. https://doi.org/10.1037/a0013110

Brooks, N., Pogue, A., & Barner, D. (2011). Piecing together numerical language: Children's use of default units in early counting and quantification. *Developmental science*, *14*(1), 44–57.

Carey, S. (2009). *The origin of concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001
Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In Anna H.-J. Do et al. (Ed.), *Proceedings of the 25th Boston University conference on language development*.
Cresswell, M. J. (1976). The semantics of degree. In *Montague grammar* (pp. 261–292). Academic Press.
Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. A. G. CAREY. Rochelle (Ed.), *The epigenesis of mind* (pp. 293-322). Erlbaum Associates.
Gelman, R., Cohen, M., & Hartnett, P. (1989). To know mathematics is to go beyond thinking that "fractions aren't numbers". In C. Maher, G. Goldin, & R. Davis. (Eds.), *Proceedings of the eleventh annual meeting of North American Chapter of the International Group for the Psychology of Mathematics Education* (pp 29-67).
Giralt, N., & Bloom, P. (2000). How special are objects? Children's reasoning about objects, parts, and holes. Psychological Science, 11(6), 497–501.
Goldin-Meadow, S., Seligman, M. E., & Gelman, R. (1976). Language in the two-year old. *Cognition*, *4*(2), 189–202. https://doi.org/10.1016/0010-0277(76)90004-4
Goswami, U. (1989). Relational complexity and the development of analogical reasoning. *Cognitive development*, *4*(3), 251–268.
Gotowski, M., & Syrett, K. (2020). Investigating the hypothesis space of children's interpretation of comparatives. Proceedings of the 44th annual Boston University. Conference on language development (pp. 154–167). Cascadilla Press.
Gualmini, A. (2004). Some knowledge children dont lack. *Linguistics, 42*(5), 957–982.
Hartnett, P., & Gelman, R. (1998). Early understandings of numbers: Paths or barriers to the construction of new understandings? *Learning and Instruction*, *8*(4), 341–374. https://doi.org/10.1016/S0959-4752(97)00026-1
Ji, Y., & Papafragou, A. (2020). Midpoints, endpoints and the cognitive structure of events. *Language, Cognition and Neuroscience*, *35*(10), 1465–1479. https://doi.org/10.1080/23273798.2020.1797839
Kennedy, C. (1999). Gradable adjectives denote measure functions, not partial functions.
Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, *30*(1), 1–45.
Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, *81*(2), 345–381.
Lenth, R. (2018). Package 'lsmeans'. *The American Statistician*, *34*(4), 216–221.
Lewis, S., Hacquard, V., & Lidz, J. (2017). "Think" pragmatically: Children's interpretation of belief reports. *Language Learning and Development*, *13*(4), 395–417. https://doi.org/10.1080/15475441.2017.1296768
Mathis, A., & Papafragou, A. (2022). Agents' goals affect construal of event endpoints. *Journal of Memory and Language*, *127*, 104373. https://doi.org/10.1016/j.jml.2022.104373
Melgoza, V., Pogue, A., & Barner, D. (2008). A broken fork in the hand is worth two in the grammar: A spatio-temporal bias in children's interpretation of quantifiers and plural nouns. In D.S. McNamara and J.G. Trafton, (Eds.), *Proceedings of the 30 th annual conference of the cognitive science society* (pp. 1581–1585).
Mix, K. S., Levine, S. C., & Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology*, *35*(1), 164. https://doi.org/10.1037/0012-1649.35.1.164
Rotstein, C., & Winter, Y. (2004). Total adjectives vs. partial adjectives: Scale structure and higherorder modifiers. *Natural Language Semantics*, *12*(3), 259–288. https://doi.org/10.1023/B:NALS.0000034517.56898.9a
Salmon, N. (1997). Wholes, parts, and numbers. *Philosophical Perspectives*, *11*(s11), 1–15. https://doi.org/10.1111/0029-4624.31.s11.1
Shipley, E. F., & Shepperson, B. (1990). Countable entities: Developmental changes. *Cognition*, *34*(2), 109–136. https://doi.org/10.1016/0010-0277(90)90041-H
Smith, C. L., Solomon, G. E., & Carey, S. (2005). Never getting to zero: Elementary school students' understanding of the infinite divisibility of number and matter. *Cognitive Psychology*, *51*(2), 101–140. https://doi.org/10.1016/j.cogpsych.2005.03.001
Sophian, C. (1997). Beyond competence: The significance of performance for conceptual development. *Cognitive Development*, *12*(3), 281–303. https://doi.org/10.1016/S0885-2014(97)90001-0
Sophian, C. (2000). Perceptions of proportionality in young children: Matching spatial ratios. *Cognition*, *75*(2), 145–170.
Spinillo, A. G., & Bryant, P. E. (1999). Proportional reasoning in young children: Part-part comparisons about continuous and discontinuous quantity. *Mathematical Cognition*, *5*(2), 181–197. https://doi.org/10.1080/135467999387298
Srinivasan, M., Chestnut, E., Li, P., & Barner, D. (2013). Sortal concepts and pragmatic inference in children's early quantification of objects. *Cognitive Psychology*, *66*(3), 302–326. https://doi.org/10.1016/j.cogpsych.2013.01.003
Syrett, K., & Aravind, A. (2022). Context sensitivity and the semantics of count nouns in the evaluation of partial objects by children and adults. *Journal of Child Language*, *49*(2), 239–265.
Syrett, K., Kennedy, C., & Lidz, J. (2010). Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics*, *27*(1), 1–35.
Wagner, L., & Carey, S. (2003). Individuation of objects and events: A developmental study. *Cognition*, *90*(2), 163–191. https://doi.org/10.1016/S0010-0277(03)00143-4