‘Five’ is the number of bunnies and hats: Children’s understanding of cardinal extension and exact number

Anonymous CogSci submission

Abstract

When do children understand that number words (such as ‘five’) refer to exact quantities and that the same number word can be used to label two sets whose items correspond 1-to-1 (e.g., if each bunny has a hat, and there are five hats, then there are five bunnies)? Two studies with English-speaking 2- to 5-year-olds revealed that children who could accurately count large sets (CP-knowers) were able to infer that sets exhibiting 1-to-1 correspondence share the same number word, but not children who could not accurately count large sets (subset-knowers). However, not all CP-knowers made this inference, suggesting that learning to construct and label large sets is a critical but insufficient step in discovering that numbers represent exact quantities. CP-knowers also failed to identify 1-to-1 corresponding sets when faced with sets that had an off-by-one difference, suggesting that children who could accurately count large sets used approximate magnitude to establish set equality, rather than 1-to-1 correspondence. These results suggest that children’s initial intuitions about numerical and set equality are based on approximation, not 1-to-1 correspondence, and that this occurs well after they have learned to count and construct large sets.

Keywords: number words; number concepts; exact equality; 1-to-1 correspondence; cardinal extension; language development

Introduction

Imagine you attend a popular conference talk where every chair is occupied. After the talk, you want to know how many people attended. Is there a way to know? As numerate adults, we know that we can count the number of chairs and infer the number of attendees. This is because we know that two sets have the same number of items if and only if their members can be placed in 1-to-1 correspondence, and also that equal sets receive the same numerical label. Reasoning about exact equality between sets is a key milestone of numerical development and in understanding exact number word meaning.

How do children acquire this knowledge? Children as young as 2 years old deploy a counting procedure by pointing to each item in a set as they recite numbers in a stable order – e.g., ‘one’, ‘two’, ‘three’ (Gelman & Gallistel, 1986), but don’t yet know how counting represents exact number. Using the Give-N task (Wynn, 1990, 1992), researchers have found that children begin by learning the meanings of ‘one’, then ‘two’ and ‘three’ (e.g., giving one object when asked for ‘one’), during which period they are called “subset-knowers” (because they know only a subset of number word meanings). Eventually, children learn to accurately give larger numbers by counting and giving all items that are implicated in their count – at which point they are often called “Cardinal Principle knowers” (“CP-knowers”) on the premise that they understand that the last number in a count represents the cardinality of the counted set.

According to one prominent hypothesis, as soon as children acquire their first exact number word meanings, they make an inductive inference that all number words – not just ‘one’, ‘two’, and ‘three’ – represent exact number. In support of this hypothesis, Sarnecka & Gelman (2004) presented subset-knowers with a set labeled with a number word (“Look, there are five frogs”), and found that when an item was added or subtracted from the set, children judged that a different number word should be used. Children also correctly reasoned that the same number word should be used when the transformation did not change the quantity, such as shaking the box containing the items. Notably, these judgments extended to number words beyond subset-knowers’ performance in a Give-N task, but not to other quantifiers such as ‘a lot’. Other studies, however, have questioned these findings, showing that children fail with highly similar tasks (Condry & Spelke, 2008; Sarnecka & Wright, 2013), and that simpler explanations that do not involve exact number knowledge can explain the data, including pragmatic inferences like the principle of contrast (Brooks, Audet, & Barner, 2013; see Izard, Streri, & Spelke, 2014 for a discussion).

According to an alternative hypothesis, children only truly understand how number words represent number when they become CP-knowers and learn to accurately count large sets. In particular, Carey (2004, 2009) proposed that learning the Cardinal Principle (CP) involves making an analogical mapping between counting and number - that for every item added to a set, one should count up one number in the count list, such that objects and number words are in 1-to-1 correspondence (see also Gentner, 2010). In support of this, Sarnecka & Carey (2008) argued that CP-knowers understand that adding one item to the set means moving forward one word in the count list (e.g., ‘four’ to ‘five’), but that subset-knowers do not. Other studies, however, have argued that not all CP-knowers succeed at this task even for small numbers like ‘four’ and ‘five’ and that many fail for larger numbers, even if they are familiar to children (Cheung, Rubenson, & Barner, 2017; Davidson, Eng, & Barner, 2012; Spaepen, Gunderson, Gibson, Goldin-Meadow, & Levine, 2018). Also compatible with a later learning trajectory, some studies found that when shown a set of, e.g., eight items, many CP-knowers are unable to exactly match this set using 1-to-1 correspondence, and instead match sets approximately (Schneider & Barner, 2020). Finally, children's success in classic Piagetian conservation tasks (Piaget, 1952) shows variability well after the age that most children become CP-knowers, suggesting that CP knowledge does not correspond with success in these tasks.

To summarize, previous studies disagree regarding the role that counting plays in discovering how number words encode exact number. Some studies argue that knowledge of exactness is learned prior to mastery of counting principles, while others argue that this knowledge emerges when children become CP-knowers, or that CP-knowledge is only a first step in understanding exact number. One approach to investigate children's ability to understand exact number meaning is testing whether children extend a number label across sets that are exactly equal – termed the 'cardinal extension' principle. Frydman & Bryant (1988) showed that when two sets were obtained after a sharing procedure, preschoolers were able to infer the cardinality of one set by counting only the other set. Muldoon, Lewis, & Freeman (2003) demonstrated that preschoolers were able to count out and construct a set that is equivalent to a set they already counted. However, these studies did not assess children's understanding of counting or the CP, leaving open the question of what role CP knowledge plays in cardinal extension.

Other studies investigating how knowledge of the CP relates to cardinal extension found that CP-knowers perform better than subset-knowers in cardinal extension tasks. For example, when shown two sets that appear equal in number, CP-knowers often correctly judge that if a number word, e.g., 'five,' applies to one set, it should also apply to the other, but subset-knowers fail to do so (Sarnecka & Gelman, 2004; Sarnecka & Wright, 2013). Nonetheless, these studies focused on reporting differences between CP-knowers and subset-knowers, but did not explore variability among CP-knowers to establish whether all CP-knowers succeed (supporting the idea that exact number knowledge emerges with acquisition of the CP) or if instead, only some CP-knowers succeed (compatible with a later learning trajectory).

Another limitation of previous studies is that they did not investigate whether children based their reasoning on exact set equality or simply noticed that two sets have approximately the same number of items. Approximate number representation is used to discriminate sets in infants as young as 6-month-olds (Xu & Spelke, 2000), but is limited by the ratio between sets. This limit decreases in older children and adults, but still inhibits their ability to accurately determine whether two sets are exactly equal (Halberda & Feigenson, 2008; Huntley-Fenner & Cannon, 2000). Meanwhile, exact symbolic number knowledge supports precise representations of numerosities, allowing for accurate set comparisons even for sets beyond the subitizable range with perceptually non-discriminable ratios (see Feigenson, Dehaene, & Spelke, 2004 for a discussion).

The present studies aimed to address the limitations of past reports. We presented subset knowers and CP-knowers with two sets of animals in unequal numbers (e.g., 5 bunnies and 7 lions). Animals in each group carried items (e.g., the bunnies had blue hats and the lions had red hats), such that the number of items was exactly equal to the number of animals in each group. One set (e.g., the bunnies) was then hidden, and children were prompted to infer how many animals were hidden. We asked whether they would count the correct visible objects (e.g., the bunnies' hats) to infer the number of hidden objects, compatible with knowledge that two sets have the same number of items - and deserve the same numerical label - if they stand in 1-to-1 correspondence. In Study 2, we provided a stronger test of whether children who succeed in reasoning about set equality do so through reasoning about exact equality or approximate magnitudes. To do so, we contrasted a condition that required reasoning about 1-to-1 correspondence vs. a condition that could be solved using approximate number knowledge.

Study 1: When does children's understanding of cardinal extension develop?

Methods

Participants A preregistration is available at [anonymized link]. Eighty children were recruited from preschools in the US and Canada, and a children's museum in the US. All participants spoke English as a primary language. We excluded two participants who missed more than one trial of the Cardinal Extension task based on preregistered criteria. Our final sample included 78 children, with 38 subset-knowers (26F, 12M; M_{age} = 3.63 [2.13; 5.20]; SD_{age} = 0.75) and 40 CP-knowers (21F, 19M; M_{age} = 4.70 [3.08; 5.95]; SD_{age} = 0.61).

Materials & Procedure All stimuli and scripts for the experiment and analysis are available at [anonymized link].

Give-N. Participants were given a titrated Give-N task (following the procedure in Wynn, 1992) to assess their understanding of the CP. Participants were shown a box with fish and a plate, and were asked to "put N fish on the plate" All participants started with 'five', received increasingly higher numbers (N+1, up to 'six') if they succeeded and lower numbers (N-1) if they failed. We recorded the highest number for which a participant can construct correct sets two out of three times. If the child constructed the wrong set, they were prompted once to "count to make sure" and allowed to fix their response. We recorded the highest number for which a participant can construct correct sets two out of three times. Children who could not construct a set of 'one' were not tested based on preregistered criteria. Children who succeeded on N=6 two out of three times were designated as CP-knowers, and those who did not were designated as subset-knowers (per criteria used in Wynn, 1992).

Cardinal Extension. Materials were prepared and presented as a slidedeck with recorded audio descriptions. Participants were first introduced to the animals used in the task (lions and bunnies). They saw two familiarization trials with three animals, and each animal was associated with one item (e.g., each bunny has a bike). They were asked to report the number of animals by pointing to the screen and counting to familiarize them with the expectations in the critical trials.

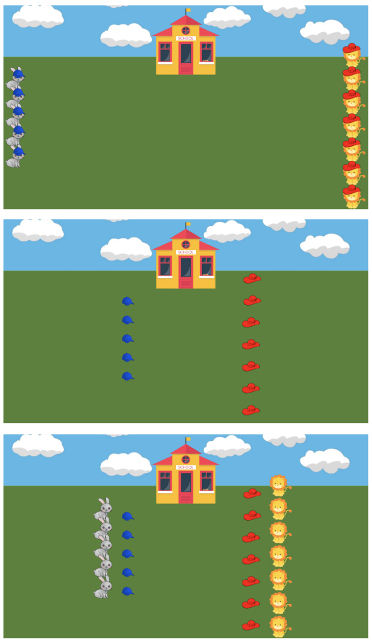

Figure 1: Schematic of Study 1, with slides on the left and the corresponding audio description on the right.

In each critical trial (Figure 1), participants saw two unequal groups of animals (e.g., 5 bunnies and 7 lions), each on one side of the screen. Each animal group was shown in 1-to-1 correspondence with an item set (e.g., 5 bunnies – 5 blue hats vs. 7 lions – 7 pink red hats). Animal and item sets were organized in a line with equal spacing to facilitate element tracking and counting. The animals then put down their items and disappeared into a building. Participants were then asked for the number of one set of animals (e.g., "How many bunnies are in the school?"). We reasoned that if children recognized that sets in 1-to-1 correspondence share the same number, they should use the correct items to infer the number of animals (e.g., counting the blue hats when asked about the bunnies). Participants were encouraged to point and explicitly count the items. If counting resulted in a different response, we analyzed the final count. Participants were allowed one opportunity to fix a wrong response.

Participants saw six critical trials in total: three small-set trials where sets < 4, and three large-set trials where sets > 4. Participants saw one out of six pseudo-randomized trial orders. Trials were counterbalanced for the target animals, side of set appearance, and which animal set is larger.

Highest Count. This task was included as a general proxy of counting experience, to allow us to differentiate between CP-knowers with different degrees of counting expertise. Participants were asked to "count as high as [they] can," beginning from one. The experimenter prompted them with "one" (but no other number) if they failed to respond, and they were prompted once after they stopped to keep counting. We recorded the highest number they reached without errors. As a preview, Highest Count was not a significant predictor of performance in either study, so we omitted results related to this measure for the purposes of this short paper.

Results & Discussion

Our primary question was whether CP-knowers were more likely to succeed at the Cardinal Extension task compared to subset-knowers across both small and large sets. We first asked whether participants selected the correct set, either by pointing to the set that was equal to the target animal set, or by giving the correct count for the target animal set[1]. Two-tailed one-sample t-tests showed that only CP-knowers (M = 0.88, SD = 0.33) performed better than chance at identifying the correct item set to infer the number of hidden animals ($t(239) = 17.98$, $p < .001$), and they succeeded for both small and large sets (ps $< .001$). Meanwhile, subset-knowers (M = 0.56, SD = 0.50) performed at chance ($t(227) = 1.73$, $p = .085$), and failed to identify the correct set for both small and large sets (ps $> .05$). Additionally, there was variability in CP-knowers' performance: out of 40 CP-knowers, only 26 (65%) succeeded in pointing out the correct set in all six trials. 4 CP-knowers (10%) succeeded in only half of the trials or fewer. These results provide evidence that the ability to reason about cardinal extension does not develop when the CP is acquired, but rather after the CP knowledge stage.

In order to further analyze the effect of CP knowledge on cardinal extension performance, we constructed generalized mixed-effects logistic regression models (GLMMs) predicting correct set selection with age (z-scored), set size (Small/Large), knower level (CP-knowers / subset-knowers), and knower level*set size interaction as fixed effects. We compared all models against a base model that included only age and set size as predictors. All models included by-subject random intercepts, but we omitted preregistered by-item random intercepts due to overfitting.

The model that best explained our data was the base model consisting of only age and set size as predictors (Figure 2A). In this model, only age predicted success in identifying the correct set, and older children were better than younger children ($\beta = 1.66$, 95% CI [1.05, 2.28], $p < .001$). Set size (small vs large sets) did not further explain variation in performance ($p = .573$). Exploratory models analyzing subset-knowers and CP-knowers separately showed that this age effect was driven by only subset-knowers ($\beta = 1.60$, 95% CI [0.79, 2.41], $p < .001$). Age did not predict performance in the CP-knower group ($p = .312$).

Adding knower level (M1) and knower level*set size inter-

[1] We assumed that children would not be able to report the correct cardinality without identifying the correct set.

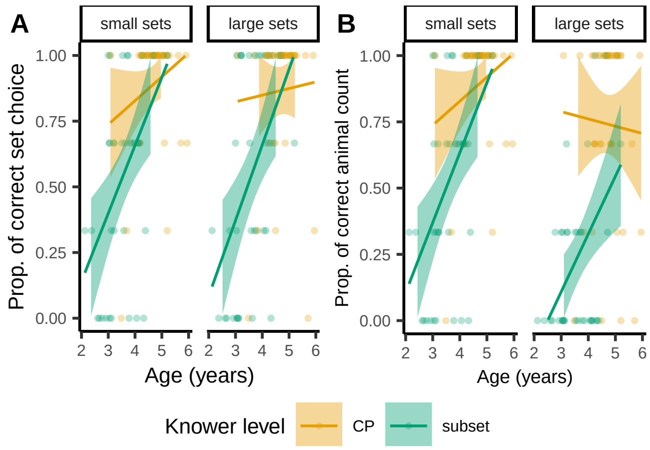

Figure 2: Each dot represents a participant. Shaded areas represent 95% confidence intervals. Responses are grouped by knower level and set size. A) Proportion of correct item set choice by participant against age. B) Proportion of correct animal quantity response by participant against age.

action (M2) to the base model did not explain significantly more variation in the data (M1: $\chi^2(1) = 3.07$, $p = .080$; M2: $\chi^2(2) = 3.28$, $p = .194$). No additional predictors were significant in either model (*p*s $> .05$). This result suggests that even though only CP-knowers succeeded at this task and subset knowers were at chance, age fully explained the difference between the two groups. We confirmed this with an exploratory model that included knower level and set size as predictors without age. Here there was an effect of knower level, with CP-knowers performing significantly better than subset-knowers ($\beta_{CP} = 2.91$, 95% *CI* [1.64, 4.17], $p < .001$). However, this model did not explain the data better than the base model with only age and set size (ΔAIC = 10.25).

In contrast to these initial analyses in which we considered success in the task as choosing the correct set corresponding to the hidden animals, in the next analysis we analyzed performance based on whether participants both choose the correct set, and correctly inferred the number of the hidden animal set by counting the correct item set. This is a more conservative measure of cardinal extension, because success requires not only attending to the equality of the animal and item sets, but also counting the selected item set accurately. We constructed another set of GLMMs with the same fixed and random effect structure as above (but also including by-item random intercepts) to predict children's success in inferring the correct number of animals.

The model that best explained the data included age, CP knowledge, and set size (Figure 2B). CP-knowers were significantly better at inferring the correct number of animals ($\beta_{CP} = 1.89$, 95% *CI* [0.67, 3.11], $p = .002$), as were older kids ($\beta = 1.04$, 95% *CI* [0.40, 1.69], $p = .001$) and in trials with small set sizes ($\beta_{small} = 1.83$, 95% *CI* [1.04, 2.62], $p < .001$). Adding knower level * set size interaction does not improve the model ($\chi^2(1) = 0.97$, $p = .326$), and the interaction was not significant ($p = .325$).

Study 2: Is CP-knowers' success in cardinal extension based on exact equality?

Study 1 showed that CP-knowers are able to reason that sets that are equal in number can be labeled by the same number word. However, it leaves open how CP-knowers might achieve this. One possibility is that they select the correct item set through noticing a 1-to-1 correspondence between items and animals (e.g., bunnies and hats). Alternatively, they might compare the approximate quantity of these sets (e.g., 'approximately seven' bunnies and 'approximately seven' hats). They might also succeed by simply noting the association between items and animals without attending to cardinality at all (e.g., the bunnies appeared with blue hats, therefore, count the blue hats).

To probe whether CP-knowers use exact or approximate quantities in reasoning about cardinal extension, and to eliminate the possibility of using identity associations between animals and items, we conducted a follow-up study that paired one animal set with two item sets in varying ratios. One item set was in 1-to-1 correspondence with the animal set. The distractor item set differed in either a perceptually discriminable manner (e.g., 5 hats - 10 bunnies), or were off by one in quantity and thus not discriminable (e.g., 9 hats - 10 bunnies). If CP-knowers succeed in cardinal extension through reasoning about 1-to-1 correspondence they should succeed in both cases. However, if they used approximate quantities, they would succeed in trials with discriminable ratios, but not in the off-by-one trials.

Methods

Participants A preregistration is available at [anonymized link]. Eighty children were recruited from preschools and a children's museum in the US. Given the failure of subset-knowers in Study 1, all participants were CP-knowers who spoke English as a primary language. We excluded two participants who missed more than one trial of the Cardinal Extension task based on preregistered criteria. We also excluded 18 trials where participants started counting before the prompt. Our final sample included 78 CP-knowers (45F, 33M; $M_{age} = 4.66$ [3.00; 5.98]; $SD_{age} = 0.73$).

Materials & Procedure All stimuli and scripts for the experiment and analysis are available at [anonymized link]. Participants were given a Give-N task and a Highest Count task following the procedure from Study 1. Only children who were CP-knowers in the Give-N task proceeded to the Cardinal Extension task.

Cardinal Extension. Materials and procedure were similar to Study 1, with any differences noted. In the familiarization phase, participants saw an additional trial with three animals (in this study, only bunnies), but only two of them had items and one of them did not. The bunny missing an item was pointed out to the participant ("This bunny doesn't have a carrot.") Participants were also asked to point and count the number of bunnies for this trial.

In each critical trial, a set of bunnies appeared at the bottom of the screen with two sets of items. One set of items (the target set) was exactly equal to the number of bunnies, and the other set (the distractor set) had fewer items. The audio description highlighted the violation of 1-to-1 correspondence between the distractor set and the target set in all conditions, and was accompanied by gestures to the bunnies that were missing items (e.g., 'Some of the bunnies don't have their hats.'). The bunnies then put down each item in each set one at a time, further emphasizing the 1-to-1 correspondence between the bunnies and the target set and the mismatch with the distractor set. Like Study 1, the bunnies then disappeared into a building, leaving their items behind. Participants were then asked for the number of bunnies. If children use 1-to-1 to reason about exact equality, they should use only the target set to infer the number of bunnies across conditions. If participants did not respond, did not overtly count, or made a counting mistake, they received prompts from the experimenter as described in Study 1.

Participants saw nine trials in total: three small-set trials where sets < 4, and six large-set trials where sets > 4. Large-set trials included three with discriminable ratios (Large-DR) where the ratio between the bunnies and the distractor item set was $>= 2$, and three with non-discriminable ratios (Large-NR) where the distractor set had one fewer item than the number of bunnies. Participants saw one out of four pseudo-randomized trial orders and item pairings. The trials were partially-counterbalanced for order and location of item sets.

Results & Discussion

Our primary question was whether CP-knowers were equally likely to choose the correct set to infer the number of bunnies across trials of different set sizes and ratios. Similar to the analysis for the previous study, we first looked at cardinal extension performance as indexed by whether the participant selected the correct set, either by pointing to the correct set or giving the correct number. Two-tailed one-sample t-tests showed that overall performance was better than chance ($t(683) = 10.84$, $p < .001$). However, only performance in small trials (M = 0.81, SD = 0.39) and large-DR trials (M = 0.71, SD = 0.46) were better than chance (small trials: $t(228) = 12.07$, $p < .001$, large-DR trials: $t(225) = 6.86$, $p < .001$). In large-NR trials (M = 0.55, SD = 0.50), CP-knowers performed at chance ($t(228) = 1.66$, $p = .099$) (Figure 3). This suggests that CP-knowers relied on approximate number representations to complete the task, but failed whenever 1-to-1 correspondence was required to differentiate sets.

To further analyze the effect of trial type on whether participants selected the correct set, we ran GLMMs predicting correct item set choice with age (z-scored) and trial type (Small/Large-DR/Large-NR) as fixed effects. We compared this model against a base model that includes only age as fixed effects. All models included by-subject and by-item random intercepts.

We found no effect of age in our base model ($p = .123$). When trial type was included as a fixed effect, we found

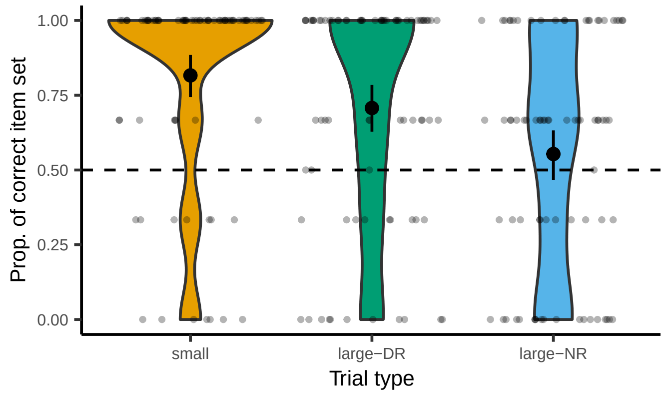

Figure 3: Proportion of correct item set choice by participant against trial type with mean proportion and 95% bootstrapped confidence intervals. Each dot represents a participant. Horizontal line indicates performance at chance=0.50.

a significant effect of trial type ($\chi^2(2)$ = 33.46, $p < .001$), and still no effect of age ($p = .127$). The enhanced model explained significantly more variation in the observed data compared to the base model in a likelihood ratio test ($\chi^2(2)$ = 15.30, $p < .001$). Post-hoc pairwise comparisons between the three trial types (Small, Large-DR, Large-NR) with Bonferroni corrections showed that CP-knowers were more likely to succeed in Small trials compared to Large-DR trials (z = 2.64, $p = .025$) and Large-NR trials (z = 5.73, $p < .001$). Success in Large-DR trials was also significantly higher than in Large-NR trials (z = 3.47, $p = .002$).

In addition to simply asking whether children chose the correct set as the basis for inferring the number of hidden bunnies, we also analyzed performance based on whether participants both inferred the correct set and counted it correctly. We constructed another set of GLMMs to predict this behavior, using the same fixed effects structure as above. All models included by-subject random intercepts, but we omitted preregistered by-item random intercepts due to overfitting.

We found that both trial type ($\chi^2(2)$ = 81.64, $p < .001$) and age ($\chi^2(1)$ = 5.64, $p = .018$) were significant predictors of whether children inferred the correct number of bunnies. Also, the model with trial type and age explained significantly more variation in the observed data compared to the base model that only included age as a predictor ($\chi^2(2)$ = 27.78, $p < .001$). Post-hoc pairwise comparisons with Bonferroni corrections showed the same pattern of success across trial types as the above analysis: CP-knowers were more successful in Small trials compared to Large-DR trials (z = 7.26, $p < .001$) and Large-NR trials (z = 8.91, $p < .001$), and success in Large-DR trials was also significantly higher than in Large-NR trials (z = 2.63, $p = .025$).

General Discussion & Future Directions

We investigated the role that counting knowledge plays in children's understanding of exact equality, and in particular that sets in 1-to-1 correspondence should be given the same numerical label. We found three main results. First, compatible with some previous studies, Study 1 found that CP-knowers were more likely than subset-knowers to infer that two sets should receive the same number label only if they are numerically equal. In fact, subset-knowers performed at chance on this task for both small and large sets. This provides evidence against the hypothesis that exact number meaning develops even before children learn the CP. Second, Study 1 found that although many CP-knowers made this inference, many also failed, suggesting that this understanding is not the product of acquiring the CP. Third, Study 2 found that although CP-knowers succeeded at cardinal extension for small sets within the subitizable range and large sets with perceptually discriminable ratios, they failed to use 1-to-1 correspondence as a cue to exact equality. This suggests that CP-knowers' initial understanding of cardinal extension is driven by sensitivity to approximate quantities, not 1-to-1 correspondence.

These results call into question the hypothesis that understanding exact number meaning occurs when the CP is acquired through a bootstrapping process where children notice an analogical mapping between counting up the count list and adding one item to a set (Carey, 2004, 2009). This account posits that children need to be sensitive to 1-to-1 correspondence in order to learn the CP, and predicts that CP-knowers should reason that numbers have exact meanings. Yet our data showed that children who know the CP performed at chance for perceptually non-discriminable ratios that are differentiated by violations of 1-to-1 correspondence. This suggests that understanding of how number words encode exact equality continues to develop well after children master the counting procedures required to become a CP-knower.

Our findings that CP-knowers fail to reason about cardinal extension in the case of large sets with perceptually non-discriminable ratios contrasts with previous results documenting CP-knowers' ability to judge that sets with off-by-one differences (e.g., 5 vs 6 items) should be denoted by different numbers (Sarnecka & Gelman, 2004; Sarnecka & Wright, 2013). Additionally, Izard et al. (2014) found that even subset-knowers use 1-to-1 correspondence between two sets to determine whether an item is missing from one set, even though this ability is not robust when sets undergo identity and quantity transformations. Considering these previous results,it is unclear why CP-knowers would depend on approximate quantities for our task. One possible explanation is that hiding the target set induced the same processing difficulties as transforming sets, thus making CP-knowers regress to judging sets based on approximate magnitudes. Alternatively, perhaps children attend to 1-to-1 correspondence in previous tasks without appreciation of its relationship to equinumerosity (see Izard et al., 2014 for a discussion). Further research is necessary to determine when and how children begin using 1-to-1 correspondence as opposed to approximate quantities to reason about set equality, and how this ability is related to other number capabilities such as the CP.

One question that arises from our results is why CP-knowers succeed at constructing large sets that are off-by-one (e.g., 'five' and 'six' fish in the Give-N task), yet failed to understand that two sets have a different number of items if they do not exhibit 1-to-1 correspondence. One possibility, proposed by Davidson et al. (2012), is that CP-knowers have acquired the ability to construct large sets as a rote procedure and therefore lack adult-like understanding of number words. For example, when asked to 'give six fish,' these CP-knowers might follow a procedure in this form: 'Begin counting from one, for each number word partition one item to a separate set, stop counting at six, and give all counted objects.' This procedure results in an accurate set of six, but requires minimal conceptual understanding of the meaning of 'six' – for example, that 'six' denotes the same cardinality across any set constructed following the procedure or that 'six' denotes 'exactly six.' Lacking this understanding, these CP-knowers might not realize that, if two sets stand in 1-to-1 correspondence, then counting one set indicates the cardinality of the second one. Another possibility is that some CP-knowers might not understand that 1-to-1 correspondence between two sets entails that they are therefore exactly equal. We note that these are potentially two different cognitive processes, and that it is an open question whether understanding that number words denote exact quantities across non-identical sets entails understanding of exact equality between sets exhibiting 1-to-1 correspondence, vice versa, or whether these two abilities are orthogonal (see Schneider, Brockbank, Feiman, & Barner, 2022 for a discussion). No previous studies have successfully differentiated these two ideas, and therefore, further research is necessary to investigate their relationship and whether one entails, or predicts, the other.

Finally, our studies found that different factors predicted children's success in choosing the correct item set whose cardinality matched that of the hidden animals vs. actually generating the correct animal count after selecting the right item set. For example, both studies showed that older CP-knowers were better than younger CP-knowers at inferring the correct number of animals, but age was not a significant predictor for CP-knowers' success in choosing the correct item set. This suggests that there might be multiple cognitive processes involved in reasoning about cardinal extension. Accurately counting a large set might be a difficult procedure, and requiring children to count in tasks that test knowledge of exact set equality and exact number meaning might underestimate their conceptual understanding.

In sum, our results provided evidence for the hypothesis that children develop exact number knowledge and reason about exact set equality in a protracted trajectory that extends after they learn the CP. Additionally, we found that CP-knowers initially reason about set equality by approximate magnitudes rather than 1-to-1 correspondence. Our findings raise questions for future research regarding when children develop understanding of exact number and exact set equality, and to what extent 1-to-1 correspondence is invoked in the process.

Acknowledgements

[Anonymized]

References

Brooks, N., Audet, J., & Barner, D. (2013). Pragmatic inference, not semantic competence, guides 3-year-olds' interpretation of unknown number words. *Developmental Psychology*, *49*(6), 1066. http://doi.org/10.1037/a0029384

Carey, S. (2004). Bootstrapping & the origin of concepts. *Daedalus*, *133*, 59–68. http://doi.org/10.1162/001152604772746701

Carey, S. (2009). Where Our Number Concepts Come From. *The Journal of Philosophy*, *106*(4), 220–254. http://doi.org/10.5840/jphil2009106418

Cheung, P., Rubenson, M., & Barner, D. (2017). To infinity and beyond: Children generalize the successor function to all possible numbers years after learning to count. *Cognitive Psychology*, *92*, 22–36. http://doi.org/10.1016/j.cogpsych.2016.11.002

Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General*, *137*(1), 22–38. http://doi.org/10.1037/0096-3445.137.1.22

Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction? *Cognition*, *123*(1), 162–173. http://doi.org/10.1016/j.cognition.2011.12.013

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314.

Frydman, O., & Bryant, P. (1988). Sharing and the understanding of number equivalence by young children. *Cognitive Development*, *3*(4), 323–339. http://doi.org/10.1016/0885-2014(88)90019-6

Gelman, R., & Gallistel, C. R. (1986). *The child's understanding of number*. Harvard University Press.

Gentner, D. (2010). Bootstrapping the Mind: Analogical Processes and Symbol Systems. *Cognitive Science*, *34*(5), 752–775. http://doi.org/10.1111/j.1551-6709.2010.01114.x

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the" number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457.

Huntley-Fenner, G., & Cannon, E. (2000). Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. *Psychological Science*, *11*(2), 147–152.

Izard, V., Streri, A., & Spelke, E. S. (2014). Toward exact number: Young children use one-to-one correspondence to measure set identity but not numerical equality. *Cognitive Psychology*, *72*, 27–53. http://doi.org/10.1016/j.cogpsych.2014.01.004

Muldoon, K., Lewis, C., & Freeman, N. H. (2003). Putting counting to work: Preschoolers' understanding of cardinal extension. *International Journal of Educational Research*, *39*(7), 695–718. http://doi.org/10.1016/j.ijer.2004.10.006

Piaget, J. (1952). When Thinking Begins. In *The Origins of Intelligence in Children*. http://doi.org/10.1037/11494-000

Sarnecka, B., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*(3), 662–674. http://doi.org/10.1016/j.cognition.2008.05.007

Sarnecka, B., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, *92*(3), 329–352. http://doi.org/10.1016/j.cognition.2003.10.001

Sarnecka, B., & Wright, C. E. (2013). The Idea of an Exact Number: Children's Understanding of Cardinality and Equinumerosity. *Cognitive Science*, *37*(8), 1493–1506. http://doi.org/10.1111/cogs.12043

Schneider, R. M., & Barner, D. (2020). Children use one-to-one correspondence to establish equality after learning to count, *CogSci*.

Schneider, R. M., Brockbank, E., Feiman, R., & Barner, D. (2022). Counting and the ontogenetic origins of exact equality. *Cognition*, *218*, 104952. http://doi.org/10.1016/j.cognition.2021.104952

Spaepen, E., Gunderson, E. A., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2018). Meaning before order: Cardinal principle knowledge predicts improvement in understanding the successor principle and exact ordering. *Cognition*, *180*, 59–81. http://doi.org/10.1016/j.cognition.2018.06.012

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193. http://doi.org/10.1016/0010-0277(90)90003-3

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

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11.