A 6-month longitudinal study on numerical estimation in preschoolers

Pierina Cheung
mcheung@wesleyan.edu
Department of Psychology
Wesleyan University

Emily Slusser
emily.slusser@sjsu.edu
Child & Adolescent Development
San Jose State University

Anna Shusterman
ashusterman@wesleyan.edu
Department of Psychology
Wesleyan University

Abstract

The current study investigated the development of numerical estimation in 3- to 5-year-old children sampled monthly for six months. At each session, children completed a task that assesses verbal number knowledge (Give-N task) and a numerical estimation task that assesses approximate number knowledge (Fast Cards). Results showed that children who acquired the cardinal principle (CP) during the course of the study showed marked improvement on the estimation task. Following CP acquisition, estimation became more accurate overall but also fluctuated widely. We discuss the implications of our findings for number word learning, particularly the mapping between verbal number and the approximate number system (ANS).

Keywords: numerical estimation; approximate number; subset-knowers; cardinal principle knowers; longitudinal

Introduction

How do children acquire number word meanings? On one view, the approximate number system (ANS) – a non-verbal representation of number – provides meanings for number words. That is, the initial meanings of number words *are* ANS representations (e.g., Dehaene, 2011; Gallistel & Gelman, 2000). The current paper investigates the mapping between ANS and verbal number in a longitudinal study, by tracking the development of number word learning and the ANS-to-word mapping in a group of preschoolers.

Starting at around the age of 2, children begin to recite the count list, but they lack meanings for number words. Slowly and gradually, over the next 1 ½ to 2 years, children learn the meanings of the first few number words in order. Using the Give-N task, Wynn (1990, 1992) showed that children first acquire the meaning of 'one' – these children give one object when asked for 'one' (1-knower). Approximately 6 months later, children become 2-knowers and have exact meanings for 'one' and 'two' but not higher numbers. Later, they acquire the meaning of 'three' (3-knower) and 'four' (4-knower). Collectively, they are called 'subset-knowers' because they have acquired a subset of number word meanings. Shortly thereafter, children understand that counting can be used to generate any set size within their count list. At this stage, children are called cardinal principle knowers (CP-knowers) because they understand that the last word of a counted set denotes the cardinality of the set and have exact number word meanings for all numbers in their count list (e.g., Gelman & Gallistel, 1978; Le Corre, Van de Walle, Brannon, & Carey, 2006).

The question of how children acquire number word meanings has sparked interest on the interface between non-verbal and verbal representations of number. In particular, some have highlighted the importance of parallel individuation – a non-verbal system for representing up to 3 or 4 individual objects in parallel – as the primary source of meanings for small number words (i.e., 'one' through 'four'; Carey, 2009; Le Corre & Carey, 2007). Nevertheless, parallel individuation has a set size limit and cannot represent exact number concepts for numbers beyond four.

Unlike parallel individuation, the ANS – a system that represents number as continuous magnitudes – has no set size limit (Dehaene, 2011; Gallistel & Gelman, 2000). The ANS follows Weber's law and exhibits scalar variability (i.e., the variability of estimates increases linearly with the mean of estimates, see Feigenson, Dehaene, & Spelke, 2004, for review). The ANS is a dedicated system for representing number: it supports numerical computations such as comparison, addition, subtraction, and multiplication (e.g., Barth, La Mont, Lipton, Dehaene, Kanwisher, & Spelke, 2006; McCrink & Wynn, 2004).

A growing body of research has examined the development of approximate number knowledge and verbal number knowledge and found conflicting results with regard to when children map number words onto representations in the ANS. In a widely cited study, Le Corre and Carey (2007) tested children's number word knowledge using Give-N and developed a task – Fast Cards – to assess the mapping of number words on ANS representations. In Fast Cards, an array of dots is presented quickly (~1s) and children are asked to estimate the number of dots without counting. Le Corre and Carey (2007) analyzed the slope of average estimates produced as a function of set size. The main interest was children's estimates beyond the capacity limit of parallel individuation – sets larger than 4. They found that many subset-knowers and some CP-knowers produced slopes that were not significantly different from zero in the large number range (> 4), suggesting that these children provided similar estimates when shown sets of 6, 8, and 10. In other words, they did not provide larger estimates for larger set sizes. Further, Le Corre and Carey (2007) found that only some CP-knowers were able to provide an estimate that approximately matched target sets; these children showed positive slopes that were significantly different from zero. Based on these findings, Le Corre and Carey (2007) hypothesized further development after the CP induction, and they set a criterion for classifying children who have mapped number words onto nonverbal numerical magnitudes as having slopes greater than or equal to 0.3. They called these children "CP-mappers". On the other hand, children were identified as "CP non-mappers" if their slopes were less than 0.3. CP non-mappers were on average 6 months younger than CP-mappers. From this cross-sectional data, they inferred that children first become CP non-mappers and eventually become CP mappers. Moreover, given that CP-knowers only mapped large sets onto the ANS *after* acquiring the cardinal principle, they argued that the ANS cannot be foundational to the acquisition of early number word meanings.

Others have, however, reported different results. Using a modified Give-N task, Wagner and Johnson (2011) found that children between the ages of 3 and 5 often *do* produce larger sets for larger requests. For example, they found that although subset-knowers gave an incorrect number when asked to give 'eight fish,' they were nevertheless more likely to generate a larger set than when asked to give 'four fish'. They took this to suggest that even subset-knowers had approximate number word knowledge, and thus, the ANS plays a crucial role in children's acquisition of number words, even before children acquire the cardinal principle. However, they did not assess 'knower-level' in the same way as other researchers in the literature, making it difficult to directly compare their results to Le Corre and Carey (2007).

More recent studies have offered some insights into these discrepant findings. First, Gunderson and colleagues (2015) showed that the range of set sizes included in the analysis for both Fast Cards and Give-N could affect the results. Specifically, the slopes of subset-knowers were found to be positive and different from zero only when the analysis included numbers that were immediately beyond the range of a child's knower-level (N+1 for an N-knower); for most children except for 4-knowers, this analysis included numbers smaller than or equal to 4. On the contrary, they did not find positive slopes when the analysis included numbers *beyond* the small number range (> 4; e.g., 'five', 'six'). These findings suggest that subset-knowers understand that numbers between N + 1 and 4 are smaller than numbers larger than 4. Given that children's representation of small numbers can be supported by parallel individuation, slopes that include numerosities in the small number range cannot provide definitive evidence for the role of ANS in number word acquisition.

Odic and colleagues (2015) offer a different view, arguing that the mapping between verbal number and the ANS is not bidirectional. In Fast Cards, children are shown an array of dots and asked to generate a numeral, and thus the mapping is from the ANS to number word (ANS-to-Word mapping); in tasks where children are asked to generate sets of objects upon a verbal request, such as Give-N, the mapping is from number word to ANS (Word-to-ANS mapping). They found that Word-to-ANS mapping develops prior to ANS-to-Word mapping, and that subset-knowers show only Word-to-ANS mappings. Given their results, it remains a possibility that ANS could play some role in children's acquisition of exact number concepts.

Taking a different approach, a recent longitudinal study tested the relation between the precision of ANS representation (Panamath task) and children's verbal number knowledge (Give-N). Shusterman and colleagues (2016) found that improvements in ANS acuity were related to the acquisition of the cardinal principle, but not to other stages in number word acquisition. They also found that the largest changes in ANS acuity were observed *at* or *after* children became CP-knowers, but not before, suggesting that while the ANS could be related to number word learning, especially cardinality, it is unlikely to be the driving force. Additionally, while the longitudinal data confirmed the knower-levels trajectory proposed by Wynn (1992), the data did not support a transition from CP non-mapper to CP mapper as proposed by Le Corre and Carey (2007).

Studies on the mapping between ANS and number words thus far suggest that the ANS and number word acquisition may be related (Odic et al., 2015; Shusterman et al., 2016; Wagner & Johnson, 2011, but see Le Corre & Carey, 2007; Gunderson et al., 2015), but none have clarified how *changes* surrounding children's number word acquisition are related to mapping to the ANS. To address this, we analyzed the numerical estimation data from the Fast Cards task in a 6-month longitudinal study, with two goals: First, we aimed to investigate children's mapping between the ANS and number words, and specifically, the relation between changes in the ANS-to-Word mapping and changes in children's number word knowledge. If mappings to ANS representations do in fact support the transition from subset to CP knower, then improvement in the quality of ANS-to-Word mappings should be evident prior to this transition. The second goal was to understand the dynamics of CP knowers' mapping between number words and ANS representations. If CP-knowers are indeed a heterogeneous group with respect to the quality of ANS-to-Word mappings (i.e., non-mappers and mappers), then individual children should exhibit some stability in their performance and clear improvement over time. Further, if ANS-to-Word mappings develop most substantially *after* the CP transition, as proposed by Le Corre and Carey (2007), such development should be apparent in this longitudinal data set.

Longitudinal Study

Methods

Participants A group of 33 children (15 male) between the ages of 36 and 55 months (M = 46.6 months) participated. Children were recruited in Central Connecticut and spoke English as their primary language. They were tested once a month for a total of six months. All children completed at least four sessions, with most participating in all six sessions.

Procedure Participants were tested individually at local preschools. Written consent was obtained from parents. In each testing session, children completed an ANS acuity task (Panamath; Halberda, Mazzocco & Feigenson, 2008), elicited counting task, Give-N (Wynn, 1990), and the numerical estimation task (Fast Cards, adapted from Le Corre

& Carey, 2007), always in that order.[1]

Elicited Counting. Children were asked to count a set of 10 objects.

Give N. Children were shown a set of 20 fish and a large bowl. The experimenter started by asking the child to put one fish in the bowl ("Can you make one fish go swimming?") If the child gave 1 fish, the experimenter asked for N + 1 fish ("Can you make two fish go swimming?") After the child placed some fish in the bowl, the experimenter asked to confirm, "Is that N?" If the child responded "no", they were given the opportunity to fix it. If the child succeeded, the experimenter moved on to the next number, and if the child failed, the experimenter asked for N – 1. This process continued until the child could give 8 fish correctly. A child was classified as an N-knower if s/he was correct 2 out of 3 times when N was asked, and failed 2 out of 3 times on requests of N + 1. Children who had only acquired a subset of the numeral meanings were collectively termed "subset-knowers" (1-knowers, 2-knowers, 3-knowers, 4-knowers). Children who could give all sets correctly up to 8 were classified as cardinal principle knowers (CP-knowers).

Fast Cards. At the beginning of the task, children were told that they would see some objects on the screen. They were encouraged to guess how many objects there were, as quickly as possible, without counting. The study began with a training session in which a set of objects, ranging from 1 to 15 objects, was presented in numerical order.

There were four blocks in the Fast Cards task. In each block, there were seven trials. Sets of 1, 2, 3, 4, 6, 10, 14 objects were used. Because it is questionable whether children could discriminate sets of 8 and 10 items, set sizes used in the original Fast Cards task, we changed the stimuli to clearly discriminable sets of 6, 10, and 14. In each trial, identical objects of the same kind were presented. Each trial lasted for 1s. Numbers were randomized within each block. On two of the blocks, total surface area co-varied with numerosity, and on the other two, total surface area was held constant.

Results

Elicited Counting. All children could count up to 10 with no more than one error.

Give-N [2] At the beginning of the study, we found 14 subset-knowers (3 1-knowers, 3 2-knowers, 4 3-knowers, 4 4-knowers), and 19 CP-knowers. At the end of the study, 8 of the subset-knowers became CP-knowers, 6 children remained subset-knowers, and 19 children remained CP-knowers throughout the study.

Fast Cards In the following analysis, we focused on three main questions. First, we aimed to document developmental changes in children's estimation ability. Second, we investigated whether the ANS is related to the cardinal principle acquisition. Third, we sought to replicate previous findings that there are two groups of CP-knowers – mappers and non-mappers. To address these questions, we analyzed the development of children's estimation ability by calculating average slopes and error rates.

Relationship between changes in children's estimation ability and verbal number knowledge To examine whether the mapping between ANS and number words differed depending on children's knower-level status, we categorized children into three knower-level groups based on developmental changes observed on the Give-N task, and included this as a variable in our analysis. The "subset-only group" included children who were subset-knowers throughout the study; the "subset-to-CP group" included children who acquired the cardinal principle during the 6-month period, and "CP-only group" included children who were CP-knowers throughout the study.

Linear slopes. To examine whether children have mapped large number words onto non-verbal number representations, we assessed the slope of children's estimates as a function of target set size. Linear slopes were computed separately in the small (1-4) and large number range (6, 10, 14). Given that our main question concerns the role of the ANS, we focused our analysis on the large number range because sets less than 4 can also be represented by parallel individuation.[3] If ANS representations support cardinality development, or vice versa, we should observe significant improvement in ANS-to-Word mappings in children who acquired the CP during the course of the study (subset-to-CP).

To test this, we asked whether children's slopes differed as a function of their Knower-Level Group (i.e., subset-only, subset-to-CP, and CP-only). We constructed mixed effects model, with random intercepts for subjects, and Age (in months), Session and Knower-Level Group as fixed effects. We found no effect of Age, $X(1) = .039$, $p = .84$, or Session, $X(5) = 4.61$, $p = .46$, suggesting that children's ANS-to-Word mapping did not significantly improve over time during the 6-month period. As predicted, we found a significant main effect of Knower-Level Group, $X(2) = 9.80$ $p = .0074$, and a significant interaction between Knower-Level Group and Session, $X(10) = 18.88$, $p = .042$.

The significant interaction indicated that the developmental trajectory in estimation differs depending on children's knower-level status. We conducted analyses separately on each of the three knower-level groups. Results showed that there was no effect of Session for the subset-only group, $X(5) = 1.45$, $p = .92$, and CP-only group, $X(5) = 2.87$, $p = .72$, but there was a significant effect of Session for subset-to-CP knowers, $X(5) = 20.25$, $p < .001$ (see Figure 1). As indicated in Figure 1, the subset-to-CP group

[1] The ANS acuity task was administered as part of a larger project and the data were reported elsewhere (Shusterman, et al., 2016).

[2] Longitudinal analyses of the Give-N task are reported in Shusterman, et al. (2016).

[3] Averaging across all sessions, children from all knower-level groups showed significantly positive slopes in the 1-4 range, $ps < .008$ (mean slopes: subset-only = 1.18, subset-to-CP = 1.10, CP-only = 1.07), indicating that children were able to estimate small sets and that they were engaged in the task.

demonstrated the largest improvement in slopes in the large number range.

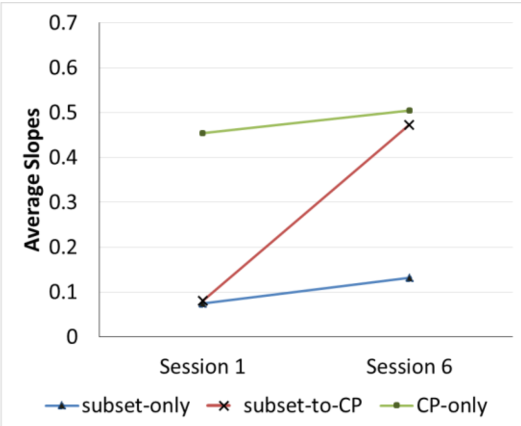

Figure 1: Average slopes from Session 1 and Session 6 in the large number range (6, 10, 14).

We further explored the changes in slopes in the large number range for the subset-to-CP group by comparing their average slopes across sessions *prior to* and *after* the CP acquisition. Consistent with the group-level analysis, we found significant improvement in slopes, $t(13.72) = -2.31$, $p = .037$. Children's slopes were close to 0 before they acquired the cardinal principle ($M_{slope} = .079$) and improved after they became a CP-knower ($M_{slope} = .33$).

While it is possible that this improvement in slopes was an effect of repeated testing, note that there was *no* effect of Session among the subset-only and CP-only groups. If practice effects are what drive children's improvements on Fast Cards, then we should observe similar levels of improvement in the subset-only and CP-only group.

Error rates. Similar analyses were conducted using error rates as a measure of accuracy. Error rates were computed as the difference between children's response and the target value. For example, when shown a set of 14 objects, the error rate was -4 if the response was 10, or +1 if the response was 15. A value of 0 indicates accurate mapping. Negative error rates indicate that children were underestimating and positive error rates indicate that they were overestimating.

Mean error rates were computed for each session. Using the same linear mixed effects model as the previous analysis on slopes, we found a main effect of Session, $X(5) = 27.22$, $p < .001$, and an effect of Knower-level Group, $X(2) = 9.54$, $p = .0085$, but there was no interaction between Knower-level Group and Session, $X(10) = 12.62$, $p = .25$. There was also no effect of Age, $X(1) = .019$, $p = .89$. Figure 2 displays average error rates for each session. Our results also showed that, consistent with previous studies, children tend to underestimate in Fast Cards (Odic et al., 2015).

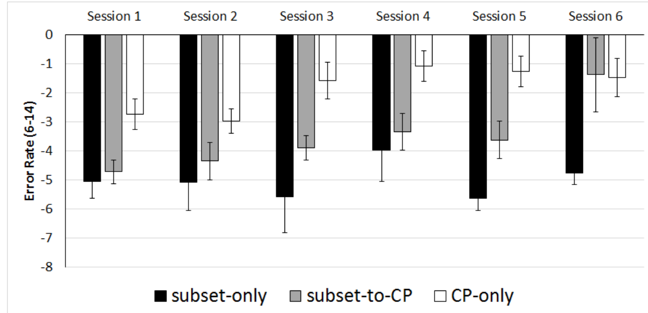

Figure 2: Average error rates across sessions in the large number range. Error bars represent SEM.

We compared overall error rates across the three groups. CP-only group was significantly more accurate than the subset-to-CP group, $t(14.69) = 2.35$, $p = .033$, and the subset-only group, $t(13.18) = 5.28$, $p < .001$. The subset-to-CP group had significantly lower error rates ($M_{error} = -3.49$) than those who remained subset-knowers throughout the study ($M_{error} = -5.17$; $t(12) = 2.26$, $p = .043$), suggesting that the ANS-to-Word mapping between subset-to-CP group and subset-only group is qualitatively different.

Results from the slope and error rate analysis both suggested that children's performance on Fast Cards improved over the course of a 6-month period. Importantly, subset-knowers who acquired the cardinal principle during the study improved to a greater extent than those who continued to be subset-knowers at the end of the study. These findings point to a previously unexplored relationship between cardinal principle acquisition and ANS-to-Word mappings—namely, that these mappings develop neither before nor after, but primarily in concert with, the CP transition. Although these results cannot definitively address the causal relation between ANS mappings and the CP acquisition, they suggest that acquiring cardinality allows children to fine-tune ANS mappings to generate more precise ANS mappings (we will return to this point in the Discussion). In our final analysis, we investigated the developmental trajectory of children's mapping ability by examining the stability of their mapper vs. non-mapper status.

Mappers vs. Non-Mappers Previous research identified separate groups of CP-knowers who had mapped large number words onto ANS as 'mappers' and those who had not as 'non-mappers' (Le Corre & Carey, 2007). The criterion for establishing mappers and non-mappers is often set at slopes ≥ 0.3, with some reporting that a criterion between 0.3 and 0.5 reveals similar results (e.g., Odic et al., 2015). To investigate whether there are two groups of CP-knowers in our sample, we conducted a Shapiro-Wilk test, which showed that CP-knowers' slopes violated the normality assumption in the first two sessions of the study (Session 1: $W = .88$, $p = .029$; Session 2: $W = .87$, $p = .001$), but the distribution was more evenly distributed from Session 3 through Session 6 ($Ws > .95$, $ps > .14$). As shown in Figure 3a, children's slopes in Session 1 had two peaks, one at around 0 and another at 0.5, but towards the end of the study in Session 6, children had a mean slope of .43 (see Figure 3b).

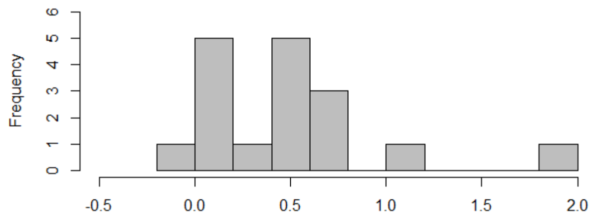

Figure 3a: Histogram of slopes in the large number range for CP-knowers at Session 1

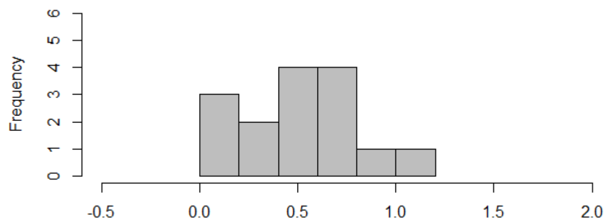

Figure 3b: Histogram of slopes in the large number range for CP-knowers at Session 6

Next, we adopted the same criterion as previous studies and identified children as 'mappers' if their slopes were ≥ 0.3 (Le Corre & Carey, 2007; Gunderson et al., 2015; Davidson et al, 2012; Odic et al., 2015). In a cross-sectional study, Le Corre and Carey (2007) found a 6-month difference in age between CP-mappers and CP-non-mappers, and Gunderson and colleagues (2015) found a 3-month difference in age. To investigate the developmental trajectory of mappers and non-mappers, we identified CP-knowers as mappers if their slopes were ≥ 0.3 at Session 1, and non-mappers if their slopes were less than 0.3. We asked (1) how likely CP-non-mappers were to become mappers at the end of the study, and (2) whether CP-mappers continued to be mappers throughout the course of the study.

We first asked how likely children who were non-mappers at Session 1 were to become mappers by Session 6. Among the 7 non-mappers, only 1 child became a mapper, 4 remained non-mappers, and 2 children did not complete Session 6. While this finding may suggest that most CP non-mappers remained non-mappers throughout the six-month period, an inspection of the individual development of slopes shows that children's slopes fluctuated over time (Figure 4). We found that the 7 non-mappers remained non-mappers for the remaining five sessions approximately 60% of the time (3 out of 5 sessions). In contrast, children who were mappers at Session 1 remained mappers 80% of the time (approximately 4 out of 5 sessions).

*Scalar variability***.** If numerals are mapped onto the ANS representations, children's estimates in Fast Cards should exhibit scalar variability (e.g., Cordes et al., 2001). That is, the mean estimates and variability should increase linearly as the set size increases. This is typically calculated as the ratio of the standard deviation over the mean estimates (i.e., the coefficient of variation, COV). We computed COVs for each target set size in each session. We hypothesized that if CP mappers used mappings to the ANS in Fast Cards, their COVs for each of the large sets should remain similar to each other.

A linear mixed effects model with Set Size (6, 10, 14), Session (1-6), children's mapper status in Session 1, and Age (in months) revealed that there was a significant effect of Set Size, $X(1) = 9.24$, $p = .002$. Pairwise comparisons revealed that COV-14 (M=.21, SE = .071) was significantly different from COV-6 (M = .32, SE = .071), p = .001, and COV-10 (M=.33, SE = .071), p = .058, but COV-6 and COV-10 did not differ from each other, p = .70. No other effects were found. Our results on COVs for 6 and 10 replicate previous studies on ANS-to-Word mapping, which did not include sets larger than 10 (Le Corre & Carey, 2007; Odic et al., 2015). Although our prediction that COVs would be constant was not validated, since the COVs for 14 were different than those for 6 and 10, we speculate that a potential explanation is that not all children could stably count to 14. Further work is needed to understand how COVs reflect children's estimation performance. Nevertheless, these results support the claim that children's estimates for sets smaller than 10 were based on mappings to ANS representations.

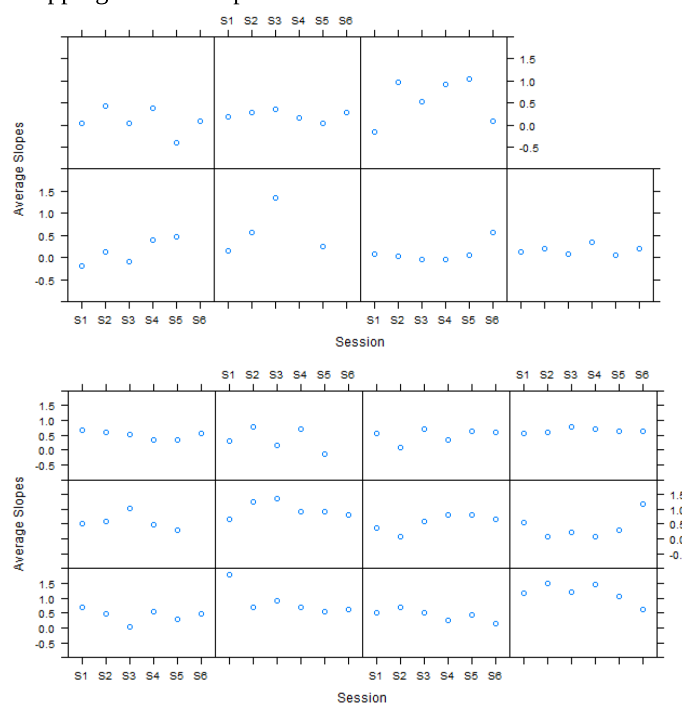

Figure 4: Scatterplots of mean slopes for each session for children who started as a) CP non-mappers (top) and b) CP mappers (bottom).

Discussion

The current study revealed two main findings. First, we found that subset-knowers who acquired the cardinal principle (subset-to-CP) showed the greatest improvement in response slopes in an estimation task – they became better at generating larger estimates for larger sets. Results with error rates also showed that these children had smaller error rates than children who remained subset-knowers throughout the study (subset-only). These results are consistent with previous studies that suggest that the ANS may be related to the acquisition of the cardinal principle (e.g., Wagner & Johnson, 2011; Odic et al., 2015; Shusterman et al., 2016).

Our second main finding is that CP knowers' estimation slopes fluctuate over time. While it is often assumed that CP non-mappers would eventually become a mapper, our results showed that this is not necessarily the case. Children's status as a mapper or a non-mapper varied from session to session, suggesting that researchers should be cautious about using 0.3 as a 'standard' cut-off point for determining mapper status. We do not suggest that the mapper and non-mapper distinction is a false dichotomy. Indeed, similar to previous studies (Le Corre & Carey, 2007; Odic et al., 2015), we also found a non-normal distribution of CP-knowers' slopes at the beginning of the study, suggesting that CP-knowers are a heterogeneous group with regard to their ANS-to-Word mapping; furthermore, children who started as mappers were relatively more likely to retain that status across sessions. Rather, our results highlight the fragile nature of children's estimation in Fast Cards, and suggest that the period following the CP transition is characterized by large fluctuation in estimation quality, rather than consistently poor estimation.

The current paper provides an important piece of data to the debate on whether the ANS drives cardinal principle acquisition. Using a longitudinal design, we found that ANS-to-Word mapping significantly improved right at the moment when children became CP-knowers, but not prior to the CP acquisition. A previous analysis from the same longitudinal study showed that the largest improvement in Weber's fraction is observed *at* or *after* children acquire the cardinal principle (Shusterman, et al., 2016). Given that the ANS-to-Word mappings fluctuate to a much greater extent than previously assumed, especially around the subset-CP transition, it seems unlikely that the ANS underlies number word acquisition. Cumulatively, the evidence suggests that while the ANS and verbal number is closely related, acquiring the CP helps children fine-tune the ANS-to-Word mappings.

If the ANS does not drive the CP acquisition, what is the precise relationship between the non-verbal and verbal number acquisition? We speculate that in acquiring the notion of cardinality, children begin to understand the exact relation between number words and sets of objects – i.e., a set has a cardinality N if and only if you follow a stable count list and are assigning each object with a number word. Through counting an exact number of objects and noticing the difference between different cardinalities, children may thus begin to more accurately estimate a set of dots in Fast Cards.

A question that this current work raises is why the slopes of children's estimates fluctuate for such a long period of time. We speculate that factors such as children's attention and processing ability may affect their verbal estimation performance. Another possibility is the length and stability of children's count list. Future studies should examine the potential of these additional factors to predict verbal estimation.

Acknowledgments

We thank all participating families and preschools. We also thank Angela Lo, Alexander Hoyle, Anthony Gutierrez, Jessica Taggart, Sarah Edelman, and Barry Finder for their help with data collection and analysis. This work was supported by an NSF CAREER grant (DRL-0845966) awarded to A.S.

References

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E.S. (2006). Non-symbolic arithmetic in adults and young children. *Cognition, 98*, 199-222.

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal and nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review, 8*(4), 698-707.

Dehaene, S. (2011). *The Number Sense: How the mind creates mathematics*. New York: Oxford University Press.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object files: Evidence from infants' manual search. *Developmental Science, 6*(5), 568-584.

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

Gelman, R. & Gallistel, C. R. (1978). *The Child's Understanding of Number*. Oxford: Harvard University Press.

Gallistel, C.R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. Trends in Cognitive Sciences, 4(2), 59-65.

Gunderson, E.A., Spaepen, E., & Levine, S.C. (2015). Approximate number word knowledge before the cardinal principle. *Journal of Experimental Child Psychology, 130*, 35-55.

Halberda, J., Mazzocco, M., & Feigenson, L. (2008). Individual differences in nonverbal number acuity predict maths achievement. *Nature*, 455, 665-668.

Le Corre, M. & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105*, 395- 438.

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, 15, 776-781.

Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition, 138*, 102-121.

Shusterman, A., Slusser, E., Odic, D., & Halberda, J. (2016). Acquisition of the cardinal principle coincides with improvement in approximate number system acuity in preschoolers. PLoS ONE, 11(4), 1-22.

Wagner, J.B., & Johnson, S.C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition, 119*, 10-22.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*(2), 155-193.

Wynn, K. (1992). Children's acquisition of number words and the counting system. *Cognitive Psychology, 24*(2), 220-251.

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