Journal of
Numerical Cognition

PsychOpen GOLD

Empirical Research

Do Preschoolers Use Rules to Represent Their Count List?

Jess Sullivan [1], Joseph Alvarez [1], Sophie Cramer-Benjamin [1], Sadie Holcomb [1,2], Melissa Nolan [1], Alex Morabito [1], David Barner [3]

[1] *Department of Psychology, Skidmore College, Saratoga Springs, NY, USA.* **[2]** *Department of Psychology, University of South Carolina, Columbia, SC, USA.* **[3]** *Department of Psychology, University of California, San Diego, La Jolla, CA, USA.*

Journal of Numerical Cognition, 2025, Vol. 11, Article e14757, https://doi.org/10.5964/jnc.14757

Received: 2024-06-05 • **Accepted:** 2025-01-27 • **Published (VoR):** 2025-04-03

Handling Editor: Lieven Verschaffel, KU Leuven, Leuven, Belgium

Corresponding Author: Jess Sullivan, Department of Psychology, Skidmore College, 815 N Broadway, Saratoga Springs NY 12866, USA. E-mail: jsulliv1@skidmore.edu

Supplementary Materials: Code, Data, Materials [see Index of Supplementary Materials]

Abstract

When children first learn to count, what do they understand about the structure of the count system? The present study investigated English-speaking children's ability to generalize the rules that structure their count list to novel contexts. A total of *N* = 86 children (3;0 – 6;11) completed a battery of tasks aimed at measuring their understanding of the English count list: they counted as high as they could, and were asked to generate successors to English numbers (e.g., "Fifty-seven: what comes next?"). Next, they were introduced to novel decade terms, and were asked to generate successors to numbers containing those terms (e.g., "Blicky-seven: what comes next?"). Children's ability to generate successors was predicted by their counting ability, and a sizeable subset of children were able to generate successors both for novel numbers and for English numbers outside their productive count range. These data suggest that emerging counters can use their understanding of the structure of the English count list to generate successors to unfamiliar numbers.

Keywords

counting, numeracy, successor function, count structure

Highlights

- 3- to 6-year-old children were tested on their counting ability and their ability to generate successors.
- Children were tested on their ability to generate successors to both familiar English number words and to novel number words.
- Children's ability to generate successors was related to their counting ability.
- Children were generally able to generate successors, even to unfamiliar and novel number terms.
- These data suggest that children may understand the structure of their count list prior to mastering the count routine.

Although children often begin to count at around age 2, they don't initially understand what number words mean, or how they are related to one another. For example, young children often begin by reciting numbers up to 5 or 10 as part of a routine, much like singing the ABCs (Fuson, 1991), although early count routines can be unstable and contain consistent inaccuracies (see Baroody & Price, 1983 for review). Crucially, these children lack an understanding of what

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0, which permits unrestricted use, distribution, and reproduction, provided the original work is properly cited.

number words mean or how they represent exact cardinalities (see Baroody & Price, 1983; Wynn, 1992). However, children ultimately learn not only the meanings of the number words in their list, but also the rules that govern the construction of the count list, including the decade+unit rule, whereby the words “one” through “nine” can be composed with decade labels like “twenty” and “thirty” to create larger numbers (e.g., Chu et al., 2020; Fuson, 1988). Understanding how to compose numbers using rules allows children to count to large numbers and to reason about numbers that they have never before encountered – such an understanding of the pattern and structure of the number system is a potentially powerful aspect of early numerical competence (Mulligan & Mitchelmore, 2009; Mulligan, Oslington, & English, 2020). In the present study, we investigated emerging counters’ understanding of how the rules of counting generalize to all numbers, including numbers that are entirely novel to them.

Substantial work has demonstrated that learning the rule structures of the number system is a critical step in early numerical competence (Moss, Bruce, & Bobis, 2015), and that children use these structures to organize their understanding of the mathematical world (Mulligan & Mitchelmore, 2009). Studies of children’s counting behaviors – which we review below – provide two specific sources of evidence that children begin to learn these rules during the preschool and early elementary school years. One source of evidence comes from children’s early counting errors. Children who have memorized their count routine should make errors that are randomly distributed across the count list, with errors becoming more likely for larger numbers and less likely for small ones. In contrast, children who use rules to combine decade labels like “*twenty*” and “*thirty*” with the units “*one*” through “*nine*” should make errors mostly at decade transitions, since in many languages (including English), decade labels like *fifty* are not transparently related to their component parts such as “*five*” or “*ten*” and therefore must be memorized. Compatible with the use of rules, English-speaking preschoolers are especially likely to make counting errors in the irregular teens (Baroody & Price, 1983; Fuson, 1988; Johnson et al., 2019; McMillan et al., 2024) and at decade transitions, often stopping at 29, 39, or 49, rather than stopping mid-decade at 24, 35, or 47 (Chu et al., 2020; Fuson, 1991; Gould, 2017; Siegler & Robinson, 1982; Wright, 1994). Further, children who stop at decade transitions can often continue counting if they are given the decade label (e.g., “actually, the number that comes after 39 is 40…can you keep counting?”; Chu et al., 2020). And, one longitudinal study of early counting showed that preschoolers had substantial knowledge of number words outside their count range – for example, one child, Daniela, could only count to 13 before making an error, but was able to recover and generate the numbers 16-25, showing knowledge of the structure of the count list beyond her range of comfortable production (McMillan et al., 2024). Taken together, these data are compatible with the hypothesis that these children have access to a rule for adding the units 1-9 to decades. Notably, these abilities are less often observed in children learning languages like Hindi and Gujarati, where it is difficult to notice the rules that govern counting because the counting systems are much less transparent and contain many exceptions to rules (Schneider et al., 2020). Meanwhile, children learning languages that are more transparent than English, like Cantonese, often count higher at earlier ages (Miller & Stigler, 1987; Miller et al., 2000; Miller et al., 1995).

A second source of evidence that children develop rule-governed knowledge of counting in the first years of life comes from their intuitions about infinity (Gelman, 1980). According to some accounts, children might come to believe that numbers are infinite by learning that numbers can be generated by rules, and that these rules can be applied iteratively, over and over, without end. Compatible with this, between the ages of 4 and 6, many children begin to claim that it’s possible to add 1 to any number, and that numbers never end (Cheung et al., 2017; Hartnett & Gelman, 1998). Further, children’s beliefs about infinity have been shown to be related to how high they can count (Cheung et al., 2017), whether they can continue counting up from decade transition errors (Chu et al., 2020), and how successful they are at generating the next number in the count list when given a prompt (e.g., “57….what comes next?”; Sullivan et al., 2023). Together, these studies suggest the emergence of a suite of beliefs compatible with a rule-governed understanding of numbers.

Data from studies of counting errors and infinity beliefs are also compatible with the idea that children use rules to govern counting prior to formal schooling (see also Schneider et al., 2021). However, two considerations limit the strength of these conclusions. First, because previous studies of children’s counting abilities are generally limited to testing relatively familiar numbers, children’s behaviors are often compatible with both rules and rote memorization. For example, the finding that Cantonese-speaking children count higher than English-speaking children (Miller & Stigler, 1987; Miller et al., 2000; Miller et al., 1995) could simply reflect the possibility that Chinese children receive more, or different, numerical training than some of their English-speaking counterparts (e.g., Miller et al., 2005). And, although the clustering of counting errors at decade transitions is compatible with the use of rules, knowledge of rules can only be inferred from these errors, because counting errors do not provide direct evidence that children can apply a rule to novel examples that fall outside their training set. Similarly, although children's beliefs about the infinity of numbers are often correlated with their counting ability (Cheung et al., 2017; Chu et al., 2020), such correlations could emerge even in the absence of the use of rules to support counting – for example, some children may simply have higher exposure both to counting *and* to ideas about infinity.

While previous studies are consistent with the idea that early counters have a rule-governed understanding of the count list, one strong piece of evidence for this would come from children's ability to apply the rules that structure their native language's count list to entirely unfamiliar contexts. The goal of the present study was therefore to test whether children could generate successors to novel numbers, and – if so – how their ability to do so relates to their understanding of counting in their native language. Specifically, we asked children to count as high as they could, and recorded their highest count prior to making an error, and their highest count when given counting feedback. This allowed us to measure their familiarity with and mastery of the English count list and their ability to recover after making counting errors. As with previous work (Cheung et al., 2017; Chu et al., 2020; Schneider et al., 2020), we reasoned these metrics would allow us to quantify not only children's familiarity with the English count list, but their ability to combine decades and units. Next, we measured children's ability to generate successors outside the context of the count routine. We presented children with familiar number words like 7 and 24 and asked them "What comes next?". Critically, our final measure was the Novel Next Number task, in which we introduced children to two novel decade terms (*mobi, blicky*), and then asked them to generate successors to numbers containing those decade terms (e.g., "what comes after *blicky-seven*?", "what comes after two hundred *mobi* six?"). We reasoned that if children's ability to count in English is driven by knowledge of rules, then children's use of rule-governed counting in English should be related to their ability to generate successors for novel number words like *blicky* and *mobi.*

In the present study, we asked four main questions: (1) Can young children generate successors for novel numbers that contain novel decade terms? (2) Do they do so as successfully as they do for English? (3) what are the predictors of their ability to generate successors to English and Novel numbers? and (4) Are their errors when generating successors consistent with the use of particular decade-unit rules for generating successors?

Method

Participants

We tested 101 participants with a pre-registered goal of 80 post-exclusion. Five children were tested but failed to meet our inclusion criteria of being native English speakers and within the age range of 3;00-6;11 years. Ten additional participants were excluded due to experimenter error (n = 4), for not having highest count data (n = 1) and for not completing enough novel number tasks (n = 5). Our final dataset included 86 participants (M_{age} = 56.9 months; Figure 1). Participants were tested at their preschools across Upstate New York region in 2019.

Materials and Procedure

Children completed three tasks: the Highest Count task, the English Next Number task, and the Novel Next Number task. Tasks were presented in the same order for all children. Study materials can be found on OSF (see Sullivan et al., 2024S). Prior to data collection, all methods and analyses were planned and written in the format of a preregistration; however, due to a miscommunication, this pre-registration was not formally submitted. A 2019 timestamped record of our preregistration is available on OSF (see Sullivan et al., 2024S). Unless specifically noted otherwise, all methods and analyses followed our original plan.

Highest Count Task

Each child was asked to count as high as they could. The highest number counted to before making an error was the child's Initial Highest Count (IHC). If the child made an error, the experimenter provided a prompt (e.g., "actually X comes after N, can you keep counting?"), and the child was permitted to continue. As in previous work (e.g., Schneider et al., 2020), the task ended when: (a) the child successfully reached 140 without prompts; (b) the child reached 140 with prompts; or (c) the child made two consecutive errors prior to reaching 140 (i.e., they could not fully recover from the prompts). Other possible ways of ending the task were planned, but never occurred in the dataset and are not discussed further.

English Next Number Task

In the English Next Number Task, the researcher prompted the child with a number and asked them to give the number that came immediately after the prompt: "In this game, I will say a number, and you tell me what comes next." This task was used to capture the child's ability to generate successors for their native count list. While rare, if the child gave a response that was n-1 instead of n+1 – indicating that they thought it was the 'what came last' game – they were reminded "the game is to find the number that comes after n, can you tell me what comes after n?". There was one training trial ("One. What comes next?") and there were 12 test trials, presented in ascending numerical order (5, 7, 16, 24, 52, 71, 105, 107, 116, 224, 252, and 271). These numbers were identical to those tested by Schneider et al. (2020), and included both numbers that we expected to be highly familiar to participants and numbers that might plausibly be well outside their productive count range.

Novel Next Number Task

As is noted in the Introduction, the purpose of the Novel Next Number task was to identify whether children could generate successors to unfamiliar decade terms. This task was structured like the Next Number Task, except that prompts contained the novel number words "Mobi" (for the Mobi block) and "Blicky" (for the Blicky block), and children were told that they were going to play a word game in the Mobi or Blicky language (order of block was counterbalanced). These tasks measured productivity for count lists that could not possibly be memorized.

For the Mobi block, children heard, "The Mobi language works like this: Five, Six, Seven, Eight, Nine, Mobi". For the Blicky block, the children heard, "The Blicky language works like this: Fifteen, Sixteen, Seventeen, Eighteen, Nineteen, Blicky." Next, the experimenter gave the participant a number in the respective 'language' and asked them to tell us what comes next – just like in the English Next Number Task (e.g., "Now, I'm going to say a number word in the Mobi Language, and you're going to tell me what comes next in the Mobi language. Ready? Mobi one"). While we initially wondered whether children might be less likely to generate successors for the Mobi block (since, in English, the numbers immediately after 10 are irregular), to foreshadow our results, children performed similarly on both blocks.

There were 14 prompts for the Mobi language, presented in the following order: Mobi 1, Mobi 5, Mobi 7, Mobi 2, Mobi 8, Mobi 4, Mobi, Mobi 6, Mobi, Mobi 3, Mobi 9 (as planned, this was not analyzed), one hundred Mobi 6, one hundred Mobi 4, two hundred Mobi 2, and two hundred Mobi 1. There were also 14 prompts for the Blicky language: Blicky 1, Blicky 5, Blicky 7, Blicky 2, Blicky 8, Blicky 4, Blicky 6, Blicky, Blicky 3, Blicky 9 (as planned, this was not analyzed), one hundred Blicky 6, two hundred Blicky 4, two hundred Blicky 2, two Hundred Blicky 1. Participants were randomly assigned to either receive the Mobi block first or to receive the Blicky block first. Seventy-nine participants completed at least some of the trials on both the Mobi and Blicky blocks, while n = 7 participants only completed one block.

Results

Data Processing for Highest Count

The Highest Count task yielded four datapoints for each child. First, the child's Initial Highest Count (IHC) was the highest number they counted to prior to making their first error. Second, the child's Final Highest Count (FHC) was the highest number they successfully counted to before the task ended. Third, we classified whether the children's first error was made at a decade transition or not. Finally, based on previous studies (Schneider et al., 2020), we classified each child as either "productive" or "unproductive". Participants were considered "productive" (n = 18) if they met one of the following criteria: (1) were able to count to 140 without error (n = 3) or (2) counted for at least 2 decades after their first error and did not make more than 3 errors during that time (n = 15). They were labeled as "unproductive" if they failed to count to 140 *even after* receiving prompts from the experimenter (n = 68). Given the relatively small number of productive counters, we do not focus on productivity classification as a primary measure in our dataset, but nevertheless report all planned analyses including this classification.

Children in our sample had a mean Initial Highest Count of 29.95 and a Final Highest of 41.93. Fewer than 10% of children in our sample could count to 100 or higher without any errors (N = 8; M_{age} = 68.3 m) and more than half of our sample (N = 45, M_{age} = 53.6 m) did not count to 20, and therefore did not count past the irregular teens. Relatively few children (N = 9, M_{age} = 59.2 m) made counting errors at decade transitions. Most of these (n = 5) could count to 29, some to 39 (n = 3), and one counted to 49 (n = 1). Instead, many children (N = 24, M_{age} = 58.4 m) made non-decade errors. Consistent with the possibility that children who make non-decade errors may have memorized their count list, 7 of these children did not count past 20.

Data Processing for Next Number Tasks

Performance on the English Next Number task and the Novel Next Number Task was coded in two ways. First, trials were coded for whether they were "correct". Correct responses contained *all* of the correct numerals and *no* incorrect numerals. For example, if the prompt was "two hundred mobi three", the child needed to say that the next number was "two hundred mobi four". They were not given credit if they said "mobi four", "two hundred four", "two hundred thirty mobi four" or "four" (all of which contain an error, despite potentially demonstrating some knowledge). Responses of "I don't know" were always counted as incorrect. However, refusal to answer was counted as missing data. Additional analyses that counted refusals to answer as 'incorrect' (as opposed to omitted) are presented in the Supplementary Materials (see Sullivan et al., 2025S) and do not differ substantially from any of the analyses reported here. Second, we coded children's responses for each place value separately. For example, if the correct answer was "two hundred mobi four", we asked whether the child provided the correct hundreds place ("two hundred"), the correct tens place ("mobi"), and the correct ones place ("four"). In this way, we could identify the location(s) of children's errors.

Finally, as planned, we first asked whether performance differed on the two Novel Next Number blocks (*mobi, blicky*). We planned to do this in order to determine whether or not data from the two Novel Next Number blocks should be combined, or analyzed separately. A paired sample t-test revealed no difference in subject-mean performance on the two blocks (t(77) = .75, p = .46). We then used regression to predict subject-mean performance on the Blicky task block mean performance on the Mobi block. This analysis revealed that the two were very highly correlated (B = 1.01, SE = .05, p < .0001, r^2 = .85). Thus, as planned, given their high correlation with one another and lack of significant difference from one another, we combined performance on the Mobi and Blicky blocks for all remaining analyses, and refer to them collectively as the "Novel Next Number Task".

Analyses

Research Questions 1 and 2: Can Young Children Generate Successors for Novel Numbers, and Do They Do so as Successfully as They Do for English?

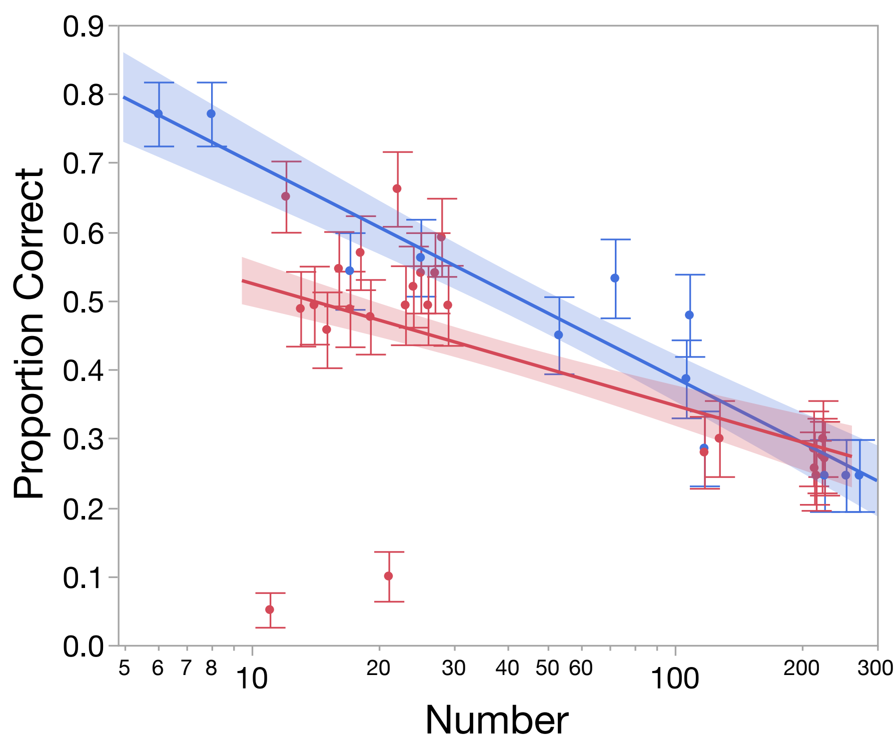
The main question of this study was whether children could generate successors for novel numbers, and if so, how their performance for novel numbers compared to their performance in English. Children generated successors at relatively high rates for both the English Next Number task (47.1% correct) and the Novel Next Number Task (42.5% correct). Performance on these tasks was highly correlated (B = .72, SE = .06, p < .0001, r^2 = .61), and did not differ significantly (t(82) = 1.98, p = .051). In a *post hoc* analysis, we asked whether there were task differences in the proportion of participants who were unable to provide *any* correct responses for the English vs. Novel Next Number tasks. Whereas only 4/84 (4.7%) of participants *never* successfully provided a successor on the English Next Number task, somewhat more failed to ever provide the correct successor on the Novel Next Number task (18/86, or 20.9%). This difference was perhaps due to a lack of rule-governed understanding of counting, but might also have been due to some children's failure to understand the requirements of the task. When considering only those children who provided at least one correct response for each task, mean accuracy was highly comparable across tasks (M_{novel} = 53.2%; M_{English} = 54.8%)

Research Question 3: What Are the Predictors of Children's Ability to Generate Successors to English and Novel Numbers?

In order to better understand the predictors of Next Number performance, we conducted planned analyses predicting binary Next Number performance (0 vs. 1) from task (Novel vs. English next number; centered), number requested (scaled), and their interaction with the following random effect structure: NNPerformance ~ Task * Scaled.Num + (Task | SubID). There was a significant effect of task (B = .87, SE = .19, p < .0001), a significant effect of numerosity (B = -1.15, SE = .08, p < .0001), and a significant interaction (B = -.56, SE = .16, p = .0003; see Figure 1). Unplanned analyses that included age in the model showed an additional significant effect of age (B = .16, SE = .02, p < .0001), and all other predictors in the model remained significant (see SOM for full reporting).

Figure 1

Proportion Correct as a Function of Number and Task



Note. Number (i.e., number requested) by task (blue = English, red = Novel). Error bars are SEM, and points represent means, even though statistical analyses were conducted on binary data. The apparent outliers for the Novel Next Number task were for when the successor to "Blicky/Mobi" was requested; these trials proved particularly challenging for children.

However, a visual inspection of the data (Figure 1) raised the possibility that this finding was driven by two trials in particular: "Mobi" and "Blicky", the trials where children were asked to provide successors to the bare novel decade term. The finding that children struggled to generate successors for novel bare decade terms, in and of itself, is of note, because it may suggest that children required the experimenter to model how the decade terms mobi and blicky could be combined with the ones place in order to successfully provide successors. Given this, we repeated the above analyses,[1] but excluded the trials where children were asked to generate successors to the bare decade terms mobi and blicky. Once again, there was a significant effect of task ($B = .73$, $SE = .23$, $p = .002$), a significant effect of numerosity ($B = -1.47$, $SE = .09$, $p < .0001$), but no significant interaction ($B = .003$, $SE = .18$, $p = .99$). When age was included in the model, it was a significant predictor ($B = .20$, $SE = .03$, $p < .0001$), and the significant interaction of task and numerosity re-emerged ($B = -.39$, $SE = .13$, $p = .004$). Taken together, these analyses suggest several things. First, performance on the English and Novel Next Number tasks was not identical when taking into account numerosity. Second, the effect of numerosity was different for the English Next Number task than for the Novel Next Number task, but this difference may have been driven in part by the trials where children were asked to provide successors to bare decade terms. Third, while age was a predictor of performance, age alone did not explain these effects. These data raise questions about the nature of the errors that children made on both tasks.

Next, as planned, we constructed models to predict mean performance for each task (English, Novel) from age (in months) and each of our counting measures (Initial Highest Count, Final Highest Count, and Productivity) resulting in six separate models. In all models, both age and the counting measure were significant predictors of successor generation (see SOM for full reporting).

Finally, we asked whether Novel Next Number Performance was predicted by English Next Number performance when taking into account counting ability. First, we predicted Novel Next Number performance from Age and English Next Number Performance; in this simple model, both English Next Number ($B = .75$, $SE = .09$, $p < .0001$) and age ($B = .007$, $SE = .003$, $p = .03$) predicted Novel Next Number performance. Next, we combined all previously significant predictors into a single model predicting Novel Next Number performance from Age, FHC, and English Next Number performance. In this model, English Next Number performance was the only significant predictor ($B = .74$, $SE = .10$, $p < .0001$). Thus, while counting ability predicted successor generation in our simple models, the strongest predictor of children's ability to generate a successor in our novel number language was their ability to generate a successor in English (see SOM for full reporting).

Research Question 4: Are Children's Errors Consistent With the Use of a Particular Decade-Unit Rule for Generating Successors?

We next explored the nature of children's errors on the Next Number task, in order to understand what types of strategies they may have deployed. When children failed to provide the correct successor, most of the time they provided an incorrect number that nevertheless followed the English conventions for composing numbers (i.e., hundreds-decades-ones: Novel = 79.7%, English = 71.4%). Some said "I don't know" (or some variation of 'I don't know'), indicating an awareness of their lack of knowledge (Novel = 12.4%; English = 25.5%), and a very small subset provided responses that contained numbers, but that were constructed in a way that failed to conform to English conventions for composing numbers (e.g., "16-Blicky-7", "sixty-forty"; Novel: 8.0%; English: 2.7%).[2] The full dataset of responses is available on OSF for individuals interested in analyzing children's errors further (see Sullivan et al., 2024S).

We next asked whether errors clustered around particular place values – e.g., when children provided the wrong successor to "one hundred mobi 5", did they make errors in the hundreds place, the tens place, the ones place, or some combination? Performance was best for the ones place, with children providing the correct digit in the ones place on more than half of trials ($52.99\%_{English}$; $54.9\%_{Novel}$). Performance did not differ for the Novel Next Number task vs. the English Next Number task in a *post hoc* paired sample *t*-test ($t(82) = 1.19$, $p = .24$; see Figure 2). Performance was similar for correctly producing the tens place ($51.0\%_{English}$; $54.4\%_{Novel}$), and once again, performance did not differ across tasks ($t(81) = 1.69$, $p = .09$). This means that children were as likely to correctly produce the unfamiliar decade terms "blicky" and "mobi" as they were to produce familiar decade terms like "twenty" and "fifty". Finally, performance for the hundreds place was substantially poorer across both tasks ($39.7\%_{English}$; $36.2\%_{Novel}$) than for the other place values, and

1) While we do not report these additional analyses in the main text, no other analyses described below are substantively impacted by the inclusion/exclusion of the "mobi" and "blicky" trials, and thus those trials are not excluded from analysis in the reporting below.

2) A tiny fraction of responses on the Novel Number task were neither numbers nor admissions of ignorance, but rather non-numerical utterances (e.g., "an animal number").

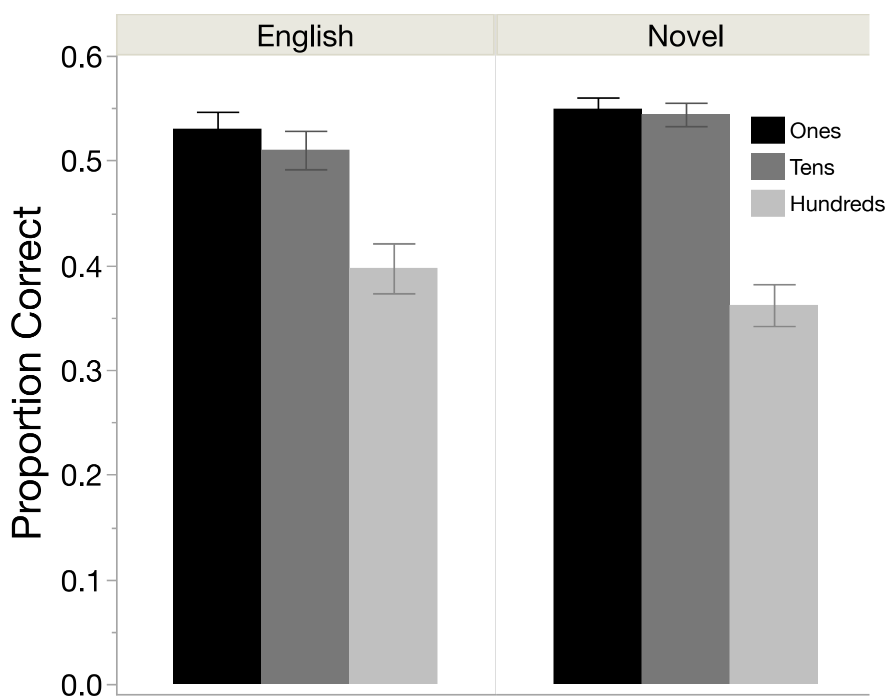
once again did not differ between tasks ($t(68) = -.61$, $p = .54$). These data suggest that children made highly similar errors across tasks, and that errors weren't restricted to a particular place value.

Did children tend to make single errors or multiple errors on a given trial? We could most easily assess this for trials that requested 2-digit responses. For the English Next Number task, among those who made errors but nevertheless provided codable responses, children generally made multiple errors (e.g., 64% of incorrect responses were incorrect because they got both the ones place and the tens place wrong while 36% made only a single error). Qualitatively, it appeared that when children made multiple errors in generating successors to English numbers, it they generated random numbers as responses (e.g., saying the successor to 53 was 20). For the Novel Next Number task, the majority of errors (59%) resulted from errors in only one place value. The remaining 41% resulted from errors across multiple place values. Qualitatively, it appeared that these multiple-errors were a mix of random-number-generation and the generation of invalid number words (e.g., "Mobi-19").

What type of errors did children make when they made an error in only one place value? A *post hoc* qualitative inspection of the data revealed that when children made errors in the tens place (but not the ones place; $n = 282$ trials), they often omitted the decade term entirely (e.g., saying that the successor to 'blicky 3' was '4' or that the successor to 116 was 107), or produced the wrong decade (e.g., saying the successor to 224 was 255). In contrast, the dominant pattern for children who made errors in the ones place (but not the tens place; $n = 302$) was to simply repeat the prompt and not generate a successor at all.

Figure 2

Proportion Correct as a Function of Place Value and Task



Note. Proportion correct for each place value for the English (left) and Novel (right) Next Number tasks. Error bars are SEM. Data are reported for the ones place (black), the tens place (dark gray) and the hundreds place (light gray), for all trials on which children provided a numerical response.

To further explore children's difficulty with the hundreds place, we generated *post hoc* descriptive statistics exploring children's performance on numbers that were within their final highest count vs. those that were outside their counting range. For the English Next Number task, children provided the correct successor 77.5% of the time for trials within their FHC range, and 30.4% of the time for numbers outside their FHC count range. For the Novel Next Number task, children produced the correct successor on 58.0% of trials within their FHC, and on 26.6% of numbers outside their highest count range.[3] These data suggest that a child's ability to generate successors was not fully determined by how high they could count, since children sometimes got both English and Novel Next Number trials within their counting range wrong, and sometimes correctly generated successors for numbers outside their counting range. However, even for the Novel Next Number Task, where memorization could not have supported performance, errors were less frequent within the child's productive count range, suggesting a role for familiarity with the English count list in shaping children's generation of successors to the novel decades *mobi* and *blicky*.

Consistent with the possibility that children used rules to generate successors to novel numbers, a *post hoc* visual inspection of the data (Figure 2) showed that performance was especially low when children had to generate successors to *mobi* and *blicky* (as opposed to generating successors for *mobi-2*). These data suggest that children were able to use both (a) the prompt provided by the experimenter and (b) their knowledge of English to construct successors to novel decade terms. When children lacked either source of information, performance decreased substantially.

Discussion

Understanding the structure of the number system is critically important for both early numeracy and for academic math success (Moss et al., 2015; Mulligan & Mitchelmore, 2009). We tested whether preschoolers could generate successors to novel numbers and whether their ability to do so was related to their knowledge of counting in their native language of English. To do this, we measured children's ability to count in English, as well as their ability to generate successors to both English number words and number words that contained novel decade terms. Children were often able to generate successors for English number words both inside and outside their counting range and also for numbers containing the novel decade terms *mobi* and *blicky*. Because children had not been introduced to the terms *blicky* and *mobi* prior to the experiment, such success could not have been due to memorization, and could only have arisen by applying the rules that structure the English count list to this novel context. In other words, their ability to generate successors to novel decade terms was likely due to their ability to use the decade+unit rule that structures the English count list (Chu et al., 2020), and their knowledge of the ordering of the unit terms in English (i.e., that 1 comes before 2 comes before 3; Fuson, 1988). Our data suggest that children are able to apply their understanding of how the English count system is structured to novel contexts, using the rules that govern English counting to generate novel successors. Given that fewer than 10% of our participants could count to 100 (and the majority of participants couldn't even count to 20 without making errors), these data suggest that the ability to generate successors for novel numbers can emerge prior to mastering the English count list.

As noted above, children were able to generate successors both for novel terms like *blicky* 7 and for numbers outside their English counting range. These data suggest that children can use their understanding of the rules for composing decades and units to generate successors for numbers far outside their counting range. The data from the English Next Number task are consistent with previous reports that have demonstrated that children have some knowledge of the number words outside of their productive count range (e.g., Barth et al., 2009; Chu et al., 2020; McMillan et al., 2024). We add to this that children are also able to generate novel numbers beyond their counting range. More generally, these data address the theoretical claim, made in past studies, that behaviors ostensibly compatible with the use of rules – such as errors in counting (e.g., Chu et al., 2020; Gould, 2017; Miller et al., 1995; Schneider et al., 2020), may indeed reflect a generalized induction regarding the structure of counting, including rules such as the decade+unit rule. These data therefore contribute to a larger literature arguing that the counting system becomes a critical organizing structure for mathematical reasoning early in development (Baroody & Price, 1983; Mulligan & Mitchelmore, 2009; Sullivan et al., 2023).

Drawing on a tradition in the counting literature of using children's counting errors to make inferences about their representations (e.g., Fuson, 1991; McMillan et al., 2024; Miller et al., 1995; Gould, 2017), we analyzed not only children's overall performance, but also the types of errors that they made. To do this, we coded children's responses for whether they provided the correct numeral in the 1s place, the 10s place, and when relevant, the 100s place. For both English

3) Based on the training component of the Novel Next Number task, we translated *blicky* as 10 and *mobi* as 20.

and the novel count systems, errors were most prevalent in the 100s place, consistent with the fact that our sample contained relatively newer counters. Indeed, given that less than 10% of our sample could spontaneously count to 100, it's perhaps more striking that children generated the correct hundreds place over 35% of the time on our next number tasks. However, for both the English and Novel Next number tasks, children's failures were equally common in the tens place and in the ones place. In other words, children's errors were not random, but rather reflected a failure to combine correct decade information with (correct) unit information. Consistent with this, performance was exceptionally low when children had to generate successors to *mobi* or *blicky* (with no ones place added). This is consistent with the idea that many preschoolers understand the rules underlying the generation of successors, but sometimes fail to correctly combine decades and units.

One of our goals was to understand the predictors of children's ability to generate successors both in English and in a novel count system. The strongest predictor of children's ability to provide successors on our Novel Next Number task was their performance on the English Next Number task, and this was true even when controlling for counting ability and age. These data suggest that while higher levels of fluency with the English count list confer a benefit for generating successors, the ability to generate successors in English provided a unique advantage for generating successors to items from a novel count list.

Our data suggest that early counters are sometimes able to generate successors to (a) novel number terms and to (b) numbers outside their productive count range. These data suggest that even some early counters may use rules to generate number words (e.g., by combining units and decades). However, children's performance for unfamiliar numbers was not perfect, they frequently made errors, and performed better for familiar numbers overall. In addition, while children were generally good at generating successors to numbers like "blicky-7", they struggled to generate successors to bare decade terms (e.g., "blicky…what comes next?"; Figure 1). These data suggest that while children may be able to combine units and decades when generating novel numbers, they benefit from being reminded how to do so -- after all, when the experimenter says something like "mobi-5, what comes next?", the experimenter has modeled one possible rule for combining decades and units.

The present data are not without their limitations. First, we only tested English-speaking children. As noted in the Introduction, English is a moderately transparent count language, meaning that it is relatively easy to identify the pattern underlying the combination of decades and units. The present data would predict cross-linguistic variability in performance on Novel Next Number tasks, such that learners of languages where the structure of the count list is more opaque (e.g., Hindi) may be slower to abstract a rule for generating successors, while learners of transparent count systems may have an easier time learning rules for generating successors. Future research should measure performance on Novel Next Number tasks across language groups in order to assess the causal role of the transparency of count list structure in shaping the ease with which children abstract rules for combining decades and units. In addition, our novel count systems were base-ten systems modeled closely on the English count system; the present data do not speak to children's general ability to generate successors across different count systems, or their ability to attach numerical meaning to the number words they generated -- a critical component of early numeracy (e.g., Mulligan & Mitchelmore, 2013). Instead, our study focused more specifically on children's ability to apply the rules used for generating successors in English to novel contexts. More generally, there is good evidence to suggest that children's performance on counting tasks depends on the nature of the task (e.g., Johnson et al., 2019), and therefore the possibility that some elements of children's performance were impacted by the nature of our task remains.

In addition, while there was a strong relation between the ability to generate successors in English and the ability to generate successors in a novel count system, the present data do not demonstrate the amount of data needed to infer the abstract rules for generating successors. Because most of our participants were newer counters with very limited mastery of the English count, it is clear that total mastery of the English count system through 100 is not required for learning some abstract rules for generating successors. That said, our current participants were neither at floor nor at ceiling, and thus future research would benefit from more explicitly mapping the relation between mastery of one's native count list and their understanding of the rules for generating successors, perhaps adopting longitudinal methods (McMillan et al., 2024).

The present study demonstrated that novice counters can successfully generate successors to novel decade terms. Their ability to do so was predicted by their mastery of the English count list, and due to the use of a novel count

PsychOpen GOLD

system, could not be attributed to memorization alone. Instead, these data provide some of the first evidence that emerging counters have access to rules for generating successors, and that these rules are robust enough to apply to novel decade terms. Children's ability to generate successors was strikingly similar for English numbers and novel numbers, consistent with the possibility that children routinely use rules to generate successors when counting (as opposed to simply memorizing the count list).

Funding: This work was supported by NSF#1749524 to J.S. and NSF #1749518 to D.B.

Acknowledgments: Special thanks to students in J.S.'s "Cognitive Development with Lab" course for assistance with data collection and data processing, to Daniel Peterson for helpful comments on the manuscript, and to Zoey Fiber for assistance with data collation.

Competing Interests: The authors have declared that no competing interests exist.

Ethics Statement: This work was approved under Skidmore College's IRB #1608-528.

Data Availability: The research data for this study are publicly available (Sullivan et al., 2024S).

Supplementary Materials

The Supplementary Materials contain the following items:

- Preregistration, tasks, research data, and code (Sullivan et al., 2024S)
- Extended reporting of analyses (Sullivan et al., 2025S)

Index of Supplementary Materials

Sullivan, J., Barner, D., Cramer-Benjamin, S., Alvarez, J., Holcomb, S., Morabito, A., Schneider, R. M., & Nolan, M. (2024S). *Novel next number* [Preregistration, tasks, research data, and code]. OSF. https://osf.io/9uwhz/

Sullivan, J., Alvarez, J., Cramer-Benjamin, S., Holcomb, S., Nolan, M., Morabito, A., & Barner, D. (2025S). *Supplementary materials to "Do preschoolers use rules to represent their count list?"* [Extended reporting of analyses]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.16148

References

Baroody, A. J., & Price, J. (1983). The development of the number-word sequence in the counting of three-year-olds. *Journal for Research in Mathematics Education, 14*(5), 361–368. https://doi.org/10.5951/jresematheduc.14.5.0361

Barth, H., Starr, A., & Sullivan, J. (2009). Children's mappings of large number words to numerosities. *Cognitive Development, 24*(3), 248–264. https://doi.org/10.1016/j.cogdev.2009.04.001

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. https://doi.org/10.1016/j.cogpsych.2016.11.002

Chu, J., Cheung, P., Schneider, R. M., Sullivan, J., & Barner, D. (2020). Counting to infinity: Does learning the syntax of the count list predict knowledge that numbers are infinite? *Cognitive Science, 44*(8), Article e12875. https://doi.org/10.1111/cogs.12875

Fuson, K. C. (1988). *Children's counting and concepts of number.* Springer.

Fuson, K. C. (1991). Children's early counting: Saying the number-word sequence, counting objects, and understanding cardinality. In K. Durkin & B. Shire (Eds.), *Language in mathematical education: Research and practice* (pp. 27–39). Open University Press.

Gelman, R. (1980). What young children know about numbers. *Educational Psychologist, 15*(1), 54–68. https://doi.org/10.1080/00461528009529216

Gould, P. (2017). Mapping the acquisition of the number word sequence in the first year of school. *Mathematics Education Research Journal, 29*(1), 93–112. https://doi.org/10.1007/s13394-017-0192-8

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

Johnson, N. C., Turrou, A. C., McMillan, B. G., Raygoza, M. C., & Franke, M. L. (2019). “Can you help me count these pennies?”: Surfacing preschoolers’ understandings of counting. *Mathematical Thinking and Learning, 21*(4), 237–264. https://doi.org/10.1080/10986065.2019.1588206

McMillan, B. G., Johnson, N. C., & Ricketts Schexnayder, J. (2024). Beyond counting accurately: A longitudinal study of preschoolers’ emerging understandings of the structure of the number sequence. *Mathematics Education Research Journal, 36*(2), 425–442. https://doi.org/10.1007/s13394-023-00453-1

Miller, K. F., Kelly, M., & Zhou, X. (2005). Learning mathematics in China and the United States: Cross-cultural insights into the nature and course of preschool mathematical development. In J. I. D. Campbell (Ed.), *The handbook of mathematical cognition* (pp. 163-178). Psychology Press.

Miller, K. F., Major, S. M., Shu, H., & Zhang, H. (2000). Ordinal knowledge: Number names and number concepts in Chinese and English. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale, 54*(2), 129–140. https://doi.org/10.1037/h0087335

Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: The role of number-naming systems. *Psychological Science, 6*(1), 56–60. https://doi.org/10.1111/j.1467-9280.1995.tb00305.x

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development, 2*(3), 279–305. https://doi.org/10.1016/S0885-2014(87)90091-8

Moss, J., Bruce, C. D., & Bobis, J. (2015). Young children’s access to powerful mathematics ideas: A review of current challenges and new developments in the early years. In J. Moss, C. D. Bruce, & J. Bobis (Eds.), *Handbook of international research in mathematics education* (3rd ed., pp. 165–202). Routledge.

Mulligan, J., & Mitchelmore, M. C. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal, 21*(2), 33–49. https://doi.org/10.1007/BF03217544

Mulligan, J. T., & Mitchelmore, M. C. (2013). Early awareness of mathematical pattern and structure. In L. D. English & J. T. Mulligan (Eds.), *Reconceptualizing early mathematics learning* (pp. 29–45). Springer.

Mulligan, J., Oslington, G., & English, L. (2020). Supporting early mathematical development through a ‘pattern and structure’ intervention program. *ZDM, 52*(4), 663–676. https://doi.org/10.1007/s11858-020-01147-9

Schneider, R. M., Sullivan, J., Guo, K., & Barner, D. (2021). What counts? Sources of knowledge in children’s acquisition of the successor function. *Child Development, 92*(4), e476–e492. https://doi.org/10.1111/cdev.13524

Schneider, R. M., Sullivan, J., Marušič, F., Žaucer, R., Biswas, P., Mišmaš, P., Plesničar, V., & Barner, D. (2020). Do children use language structure to discover the recursive rules of counting? *Cognitive Psychology, 117*, Article 101263. https://doi.org/10.1016/j.cogpsych.2019.101263

Siegler, R. S., & Robinson, M. (1982). The development of numerical understandings. In *Advances in Child Development and Behavior* (Vol. 16, pp. 241-312). https://doi.org/10.1016/S0065-2407(08)60072-5

Sullivan, J., Cramer-Benjamin, S., Alvarez, J., & Barner, D. (2023). Everything is infinite: Children’s beliefs about endless space, time, and number. *Open Mind: Discoveries in Cognitive Science, 7*, 715–731. https://doi.org/10.1162/opmi_a_00104

Wright, R. J. (1994). A study of the numerical development of 5-year-olds and 6-year-olds. *Educational Studies in Mathematics, 26*, 25–44. https://doi.org/10.1007/BF01273299

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

Journal of Numerical Cognition (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.

PsychOpen GOLD