Cognitive Science

Counting to Infinity: Does learning the syntax of the count list predict knowledge that numbers are infinite?

--Manuscript Draft--

Manuscript Number:	19-219R1
Full Title:	Counting to Infinity: Does learning the syntax of the count list predict knowledge that numbers are infinite?
Article Type:	Regular Article
Keywords:	Count list; Infinity; Conceptual change; Successor function; Highest count; Decade+Unit rule
Corresponding Author:	Junyi Chu Massachusetts Institute of Technology Cambridge, MA UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Massachusetts Institute of Technology
Corresponding Author's Secondary Institution:	
First Author:	Junyi Chu
First Author Secondary Information:	
Order of Authors:	Junyi Chu
	Pierina Cheung
	Rose M Schneider
	Jessica Sullivan
	David Barner
Order of Authors Secondary Information:	
Abstract:	By around the age of 5½, many children in the US judge that numbers never end, and that it is always possible to add +1 to a set. These same children also generally perform well when asked to label the quantity of a set after 1 object is added (e.g., judging that a set labeled "five" should now be "six"). These findings suggest that children have implicit knowledge of the "successor function": every natural number, *n*, has a successor, *n* +1. Here, we explored how children discover this recursive function, and whether it might be related to discovering productive morphological rules that govern language-specific counting routines (e.g., the rules in English that represent base 10 structure). We tested 4- and 5-year-old children's knowledge of counting with three tasks, which we then related to (1) children's belief that 1 can always be added to any number (the successor function), and (2) their belief that numbers never end (infinity). Children who exhibited knowledge of a productive counting rule were significantly more likely to believe that numbers are infinite (i.e., there is no largest number), though such counting knowledge wasn't directly linked to knowledge of the successor function, *per se*. Also, our findings suggest that children as young as four years of age are able to implement rules defined over their verbal count list to generate number words beyond their spontaneous counting range, an insight which may support reasoning over their acquired verbal count sequence to infer that numbers never end.

August 5, 2019

Dear Editors:

We are delighted to submit our manuscript titled “Counting to Infinity: Does learning the syntax of the count list predict knowledge that numbers are infinxite?” to *Cognitive Science* as a regular article. Despite experiencing only particular number words, children eventually believe that numbers go on forever. Here we report the first investigation into how children may discover the infinite nature of numbers by noticing productive morphological rules that generate successive number words.

We measured 4-5-year-old children’s familiarity with the count list and separately, their ability to use the decade+unit rule to generate numbers outside their familiar count list. Critically, **children’s knowledge that numbers never end depended not on the length of their familiar count list, but instead, was predicted by their knowledge of this productive decade rule.** These results suggest that children’s conceptions of number are influenced by their linguistic knowledge of number words: Upon knowing that the numbers one through nine are repeated iteratively across decades, children not only learn to generate successive number words given some starting point, but also gain access to the insight that one can *always* continue counting, and thus, that numbers never end.

The interaction between numerical reasoning, word learning, and concept acquisition is a topic of great interest in fields ranging from experimental psychology and language development, to math education and computational models of learning. Our results, suggesting that syntactic knowledge may lead human learners to discover that the natural numbers form an infinite class, should be of interest to researchers in all these fields. The results also suggest the important role that language learning plays in the acquisition of abstract concepts such as the successor relation and infinity. We believe that the originality, broad impact, and accessibility of these findings make *Cognitive Science* an appropriate venue for this work; we hope you will agree.

Given their expertise in number learning, language acquisition, and conceptual development, we would like to suggest as reviewers Elizabeth Gunderson, Steven Piantadosi, Barbara Sarnecka, Hilary Barth, Darko Odic, and Daniel Ansari. Due to a potential conflict of interest, we ask that Charles Yang not serve as a reviewer here. All authors have approved the final version for submission and the manuscript has not been previously published nor submitted to another journal for publication. The experimental work was approved by our Institutional Review Board and conducted with informed consent of parents. If you have any questions about the manuscript, Junyi Chu will be serving as the corresponding author.

Please let us know if there is any additional information we can provide. Thank you for your time and attention to the manuscript.

Sincerely,
Junyi Chu, Pierina Cheung, Rose M. Schneider, Jessica Sullivan, and David Barner.

March 3rd, 2020

Dear Dr. McNeil,

Thanks to you and the three reviewers for your helpful comments on this manuscript. We have made adjustments to the paper to respond to each of these comments, which we detail below. In particular, we have now done more to describe how we chose our measures (and why), and how we converged upon the cutoffs we used. In brief, we used the infinity interview because it allowed us to ask questions that were general in nature, and because it has been shown to correlate well with other important measures in past studies. Cutoffs (e.g., for defining productivity) were adopted from other studies in our lab that used the same criteria. We note this, and also explain why we believe these criteria are sensible, independent of their use in previous studies.

We have also incorporated your comments. We believe that pattern recognition of some form is almost certainly involved in learning morphological paradigms like the count list, and that this is perhaps a difference of terminology across fields, more than a difference in the substance of our claim. Learning a morphological paradigm involves detecting regularities amidst exceptions, much like domain-general pattern recognition. We have now made this connection in the paper.

We believe that these modifications make the logic of the paper stronger and thank you again for your comments.

Regards,

Junyi Chu

cc:
Rose Schneider
Pierina Cheung
Jessica Sullivan
David Barner

Overall list of changes

Location	Change	Suggestion from:
Introduction		
p.7	Added citations (Kaminsky, 2015; Sella, 2020) to the discussion that full successor function understanding comes after CP transition	Editor
p.8	Clarified definition of successor function	Reviewers 1 & 2
p.10	Added Table 1 to clarify what it means to learn the Decade+Unit Rule, and explained the logic of our hypothesis that noticing recursive syntax can lead children to infer numbers do not end	Reviewers 1 & 2
p.11	Added text explaining that children may acquire a successor function for generating numbers without appreciating the logical entailment of infinity	Reviewer 2
p.12	Replaced “understanding how this recursive rule relates to cardinality” with “.. relates to the possibility of adding 1”	Reviewer 2
p.14	Clarified logic of the current study, especially how measures were chosen, and clarified differences between Next Number and Highest Count as productivity measures	Reviewer 1
Methods		
p.15 (Sec 2.1)	Specified participant demographic pool	Editor
p.15, 17 (Sec 2.2.1)	Added flowchart of productivity decision tree (new Fig. 1) and referring text	Reviewer 3
p.18 (Sec 2.2.1)	Added explanations for error cutoffs in determining Productivity	Reviewers 2 & 3
Results		
p.21, 25	Renamed Section 3.1.1 and 3.1.2 headers for clarity	Reviewer 3
p.	Updated Figure 2 (now Figure 3) to better visualize errors and show participant IDs	Reviewers 2 & 3
p.25 (Sec 3.1.2)	Added text explaining analysis strategy for improved clarity	Reviewer 3
p.23 (Sec 3.1.2)	Replaced “many Productive Counters (n=9)” to “some Productive Counters (n=9).	Reviewer 3
p.26-27, 32-35. (Sec.3.1.2, 3.3)	Added Productivity by Initial Highest Count interaction to the analysis of Next Number accuracy and Infinity	Reviewer 2
p. 31 (Sec. 3.2)	Added text clarifying comparisons between our work and Cheung et al (2017).	Reviewer 1
p.32 (Sec. 3.2)	Referred to regression tables (Tables 3-5) earlier in the text for improved narrative	Reviewer 3
Discussion		

p. 39	Clarified how children might think that successor function is bounded to a finite list	Reviewer 2
p.31		Reviewer 2
p. 42-43	Added discussion of study limitations and potential alternative explanations for productive counting, such as individual differences in domain-general cognitive abilities	Editor
p. 43	Added reference to similarities in counting progressions observed in the educational literature (Frye et al, 2013)	Editor
Supplemental Materials		
online	Added materials, data, and analysis code to the Open Science Framework: https://osf.io/z6ky3/	Reviewers 2 & 3
p.13-18 (Sec. D)	Added analyses for Next Number performance and Infinity regressions using stricter productivity classification (allow only 1 error, not 3), and referring text in main manuscript (p.22)	Reviewer 2

Editor's comments

Thank you for submitting your paper for consideration at Cognitive Science. I have now received three reviews of your manuscript from consultants who are highly knowledgeable in the area of early number learning. I have also read it myself. We all agree that the study will be of interest to a general cognitive science audience. However, we do have some concerns and suggestions for improvements (appended below). Although I cannot accept the current version of the manuscript in Cognitive Science as is, I invite you to revise and resubmit based on the feedback here. I can't be certain that you will be able to address the concerns in a revision, but I am optimistic.

The reviewers' comments are appended in full below this letter. Please address each point made by the reviewers, especially Reviewer 3's point about your method of choosing various error cutoffs and Reviewer 1 and 2's concerns about the appropriateness of the explicit questioning method for determining understanding of the two aspects of infinity knowledge. Regarding Reviewer 1's point, I would welcome various robustness checks in the supplementary materials. For example, what do the result look like if you only allow for one error instead of three?

Thank you. We have now addressed these questions in our response below.

In addition, please address these points:

1. Provide any additional demographic information on the participants in your sample. Do you have information on race/ethnicity, family income, parent education? I ask because on page 4 you describe the breakthrough of

understanding number words happening around age 3.5 or 4 (in US English-speaking groups), but that's only true for middle-income children in the US. Was the current sample of children predominantly from middle-income families?

We have now provided general characteristics of the population from which we sampled in Section 2.1 (p. 13).

2. Because Cognitive Science connects to several disciplines, including both psychology and education, it would be ideal if you could tie this work more to some of the early childhood education literature. For example, early childhood educators have charted what they see as a developmental progression in early number knowledge (see the IES Practice Guide on Teaching Math to Young Children), and that developmental progression places "number after" knowledge two steps before +1 knowledge. Could these ideas be acknowledged and incorporated into the lit review or discussion?

Thank you for the suggestion; we now incorporate these ideas into the discussion (p. 41).

3. Jennifer Kaminski has a highly relevant CogSci Proceedings paper that should be cited and discussed when presenting the previous work that has shown that successor function knowledge is acquired after CP-knowledge.
https://mindmodeling.org/cogsci2015/papers/0186/paper0186.pdf

Thank you for the citation, we have now incorporated it into our introduction.

4. On page 9, you suggest two possibilities for the relation between how high a child can count and successor knowledge. However, there is also a third possibility. It's possible that children's non-numerical cognitive abilities such as pattern detection (apart from counting) could explain both. For example, children who are better at detecting patterns may more quickly become high counters and they also may be faster to develop successor knowledge. How does this possibility affect what can be concluded based on the current results? Perhaps a discussion of this issue would be helpful for readers.

We have now included a discussion of domain-general factors and how they likely impact performance. We believe that learning to extract rules from data is a type of pattern detection, and would expect a domain-general test of pattern learning to correlate with counting ability. Also important would be tests of domain-general capacities that don't relate directly to rule learning in content, but that might constrain learning - e.g., working memory. We have tested working memory capacity in another more recent study, and find that in that data set counting productivity measures have effects above and beyond working memory. We now incorporate these ideas into the discussion.

It's important to note that domain-general differences between children likely can't explain the pattern of findings we report as indicated by within-subjects comparisons. Our results point to dissociations between Highest Count and Productivity in predicting both knowledge of specific successors(on the Next Number task) and general successor function knowledge (on the Infinity interview). When controlling for Highest Count and age, Productivity uniquely predicted both children's ability to identify successors for numbers outside their rote counting range, and children's belief that numbers never end. Thus, at least some individual differences in conceptual knowledge about the successor function cannot be explained by the highest count measure. Consequently, we expect that any correlated non-numerical cognitive abilities would also not fully explain successor knowledge.

5. An additional limitation that could be discussed further is that children could possibly be labeled as productive counters here without having knowledge of a productive counting rule. Some children store scripts in memory and may use those to succeed. Would it be possible for a child who relies heavily on scripts to be coded as a productive counter on your task? In other words, it's possible that stating/reminding a child of the decade could activate memory of a verbatim script stored in memory, so they could start succeeding with that decade and the subsequent decade even though they don't really understand the full rule. Thus, the present coding system may overestimate the number of children who truly have a productive rule. Perhaps creating a stricter criteria for coding children as productive counters would lead to stronger relations between that category and infinity knowledge?

This is an interesting suggestion. In the paper we raise this possibility, and argue that it is unlikely given previously reported data and our data. On the view that children rely on rote memory, counting errors should occur at random locations in the count list, or, alternatively, should be more likely for less frequent number words than for more frequent numbers. Relying on scripts shouldn't predict particular difficulty with decade labels, because they are, if anything, more frequent than words that come before or after them (e.g., "sixty" is more frequent than either "fifty-nine" or "sixty-one"). Therefore, we shouldn't predict the attested finding that children often stop on decades (at 29, 39, 49, etc.) if they are simply relying on a memorized script. Also, assuming that children cannot recall "sixty" because they relied on a stored script, prompting of "sixty" should not be helpful for recalling "sixty-one" (not any more than prompting of "fifty-eight" should help with "fifty-nine"). Also, note that in our study, the Next Number task provides an alternative and potentially stricter test of Decade Productivity, as it does not have a rote counting component. We reasoned that if participants rely on memorized lists to succeed at the highest count task, then they should perform worse on the Next Number task for items outside their Initial Highest Count range than within it, since that is their best-memorized counting range. This was not the case (Section 3.1.2). Instead, as a group, children labeled as Productive Counters were equally successful for numbers within and outside their rote counting range. This suggests that children classified as Productive Counters did indeed have access to a productive rule for generating successors, as opposed to relying on stored sequences.

Perhaps a stricter test of productivity knowledge might involve unfamiliar or made-up numbers. For instance, a child should know that the successor of "a billion one" is "a billion two" and that the successor of "daxy-five" is "daxy-six". We now include a summary of recent work testing this hypothesis in the Discussion.

Reviewer #1:

This manuscript reports one experiment designed to examine 3- to 6-year old children's explicit understanding of two important aspects of early number knowledge - the successor function and infinity. This is an important study because it addresses aspects of number knowledge that, in my opinion, are critical to the development of the number concept, but which have been understudied, and when they have been studied, often conflated with other aspects of number knowledge (e.g., cardinality). Indeed, most previous studies have focused on cardinal knowledge and acquisition of the cardinal principle as the critical transition point from an immature to a mature concept of number, and authors like Sue Carey have even claimed that acquiring the cardinal principle marks a conceptual shift whereby children move beyond their foundational, but limited, nonverbal representations to induce the structure of the natural number system. While often mentioned in published work, the successor function has not often been examined with empirical tests. Barbara Sarnecka's work (reviewed by the present authors) stands as an exception. Nonetheless, I feel like much more work is needed before we understand the developmental of this critical aspect of number knowledge and the role it plays in children's understanding of the natural numbers. The present study therefore makes an important contribution to the literature.

This manuscript is well-written and the introduction frames the work clearly. The sample is quite large, the appropriate ages are targeted. The analyses are appropriate. I feel that the authors were careful with their conclusions and that they did a nice job drawing conclusions based on the findings even though the data was not in line with their expectations.

My only real concern about the work is what seem to me potentially inaccurate definitions of the main constructs. Their measure for infinity seems pretty sound - the infinity interview questions 1-3 target the right construct, which is whether think the numbers ever end. Questions 4-6, however, do not seem to be getting at the successor function, but moreso other aspects of infinity. That is, understanding the successor function requires a child to understand (either implicitly or explicitly) that each number, n, has a unique successor that is generated by adding one to n. Asking children whether you can make a "biggest number" any bigger by adding to it (question 4), or whether you can "add one to the biggest number" (question 5), do not capture the notion of unique successor. Likewise, even question 6, whether it's possible to "keep adding one" doesn't really get at succession. Instead, it seems to be targeting infinity. I do believe that the "next number" task, though a strong test, is a good measure of children's successor knowledge because in order to succeed in that task kids must (implicitly or explicitly) understand how to generate the next number from the number given. Merely knowing that you can keep adding on to a number to make it bigger doesn't seem to me to capture the essence of the successor function. Indeed, the present data was equivocal about the relationship between highest count (and therefore decade productivity and by proxy understanding of the

syntax of verbal counting) and successor knowledge based on their measure. I would like to see the authors make a convincing case of how their measure indexes true understanding of the successor function, rather than indexing just a different aspect of infinity.

Thank you for these comments. In our revision, we are now more explicit regarding how we define these constructs and why. Mathematically, the successor function is actually quite narrow - it simply states that for any number, N, there exists an immediate successor. On this definition, a child who "knows" the successor function would only need to know that every number is followed by another number, without necessarily knowing the names of particular numbers. For example, an adult might know that an indefinitely large number like "a zillion" has a successor despite not knowing how to label it (or what the labels for much larger numbers are).

In order to deploy successor function knowledge in the Next Number task, the child needs to know specific number labels. But technically, a child might be able to label the next number for a finite list of words, but not know the successor function - i.e., that *every* number has a successor. For this reason it is important to ask children more general questions that are not tied to specific numbers, as we have in the Infinity Interview.

We have now explained this more explicitly in the paper (p. 8) and have cited various sources that define the successor function in the way we describe (e.g., Decock, 2008; Wright, 1983).

The only other comment I had was whether asking children to explicitly discuss infinity and successor knowledge is really a good way to measure their understanding. I feel some adults might have trouble talking about the successor function, even though they clearly understand counting and how to generate the next number in a sequence. I guess I'm just wondering what the authors think is critical about asking children to demonstrate explicit knowledge of the successor function and infinity, rather than trying to devise an implicit measure (that differs from and improves upon the unit task, with it's limitations)?

Thank you for the comment. In response to this comment, we have now added text to our Introduction and General Discussion.

We agree that this is the trickiest question raised by this work, and is a general problem for psychologists - how to assess construct validity and choose between alternative measures. And we hope that as this literature proceeds, researchers will derive ever-better measures.

The question raised by the reviewer ultimately boils down to two questions: (1) whether there exist better, alternative, measures that we know to have better validity, and (2) whether it is important that knowledge be probed using implicit vs. explicit measures. Relating to the first question, we reviewed the literature on infinity and successor function knowledge and selected this battery because it has been used repeatedly with similar results in multiple past studies, and because there are no better alternatives that we know of, or that we could devise. Note that all of this work finds that children reach ceiling on our measures by age 6 or 7 (e.g., Evans, 1983; Evans & Gelman, 1982; Hartnett & Gelman, 1998; Cheung et al 2017), suggesting that adults would have no difficulty with them. Also, these measures appear to be correlated with other numerically relevant tasks. Still, because we like to have converging evidence, we coupled the infinity battery with tests of how children reason about attested numbers, like the Next Number task. We do hope that other, better, measures might emerge, but currently there are few alternatives, and those that exist find much later competence, probably because they require much more sophisticated domain-general reasoning. For example, in Falk's 2004 study, he creates a game in which the winner is the person who is able to name the largest number, and the dependent measure is whether the child chooses to go first, or to go second in the game; the inference is that children will always choose to go second if they understand that numbers can be infinitely generated, since the person who goes second can always use a rule to generate a larger number). Of course, there are many other demands of this task besides whether or not the child understands that numbers can be infinitely generated; for this reason, we elected to use the infinity battery presented in the present paper.

Regarding the question of explicitness, although we tested children with an interview that asks children to explicitly reason about numbers, we do not believe that answering such questions requires them to have explicit knowledge of the Peano axioms, per se. These axioms took humans thousands of years to explicitly derive after we first began using numerical symbols. However, the notion of infinity is much older in human history, and appears to be a basic intuition regarding the behavior of numbers that even our children seem to grasp: Success on the Infinity Interview does not require participants to spell out the successor function in terms of formal definitions like the Peano-Dedekind axioms, which we agree would be difficult for some adults.

Given that the development of these Infinity Interview responses is developmentally predictable across multiple studies, is related to other numeracy outcomes, and is direct in probing the underlying construct of interest, we believe that it is currently the best measure available, though we hold out hope that future studies will devise ever-better measures. We have now explained and justified our choice of tasks in the Introduction, so that future readers better understand our choice.

Again, I think this work is interesting and important, I just wasn't sure if the piece I think is most critical - the successor function - was adequately measured. In light of this, I feel like the fact that highest count and next number performance were related suggests that implicit knowledge may be related to being a productive counter, but that this relationship just doesn't hold for explicit successor knowledge.

Reviewer #2:

This study tests whether children's counting skills inform their understanding about infinity. Children performed two counting-related tasks and were questioned about their explicit knowledge about infinity. The results show that children who can count *productively* were more likely to believe that numbers are endless. The experiment was done in a thoughtful way and the paper is well written. Nevertheless, there are several major issues that question the significance and the validity of the findings.

There are many sentences that are difficult to make sense, difficult to understand, and even contradictory. It says "by learning that number words can be decomposed into decades and ones, children may realize that such rules can generate an infinite set of number words" (page 11). The rationale behind this logic is unclear. The highest count task tests whether children can generate the sequence of ones in the right order in any decade (note that the decades were prompted) until 99. That is, being able to correctly list from 1 to 9 after a given decade word is sufficient for a child to be categorized as a productive counter. How does that knowledge precisely let them realize the infinite nature of numbers? As an example, imagine that children are learning the sequence that starts with A1, A2, … A9, then B1, B2, … B9, all the way until Z1, Z2, … Z9 (i.e., it's a finite set). Even though children may master the whole list, this is a finite set and there is no pressure for children to think that the list must go on. Again, it's unclear how the authors have come to hypothesize that such kind of bounded sequential knowledge triggers the concept of infinity. A clearer logic behind this hypothesis is needed.

The intuition behind our hypothesis is that when children acquire morphologically complex words (e.g., walk-ed; chair-s) they can either learn words item-by-item or they can acquire rules that permit the generation of words (e.g., past tense; plural), which in theory can generate an unbounded number of forms. Children appear to favor the adoption of rules when the number of cases that can be described by a rule is large - e.g., an open class - as in the cases of the past tense and plural. Our hypothesis is that after having learned 2-3 decades in an item-based way, children derive a rule to describe the combination of "decade+unit", and this rule generates novel cases, effectively guaranteeing that numbers can be generated indefinitely (and thus that they are infinite).

We have now described this in further detail in the Introduction (p. 8-12).

It is difficult to understand the interpretation that "the [successor] function is bounded to a finite list" (p.36). In fact, this is contradictory, considering that the paper is centered on the Peano-Dedekind Axioms (1), according to which the successor function "generates an infinite set" (p.6). Such an expression should be removed.

Thank you for the comment. We agree with the reviewer that the successor function does generate an infinite set. We have edited the sentence to clarify that we are discussing children's developing knowledge as opposed to the formal successor function.

The way that productive counters are categorized seem ad hoc and is complicated. There seems to be no rationale behind certain decisions. For example, why allow 3 errors? Why up to 2 decades?

Thank you for the comment; we have now explained the choice of error cut-offs in greater detail (Methods, p.17), as we realize it does seem arbitrary without further context. While the application of 3-error/2-decade productivity criteria was not specifically designed with this study in mind, it was drawn from previously published, pre-registered reports in which this measure was shown to be significantly related to other measures of counting productivity (Schneider et al, 2020[1]). The 3 error criterion was chosen to allow for errors at decade transitions, which are often irregular, as well as 1 additional error per decade. The 2 decade criterion was chosen both to provide evidence that children's counting after an error is rule-governed, which is more difficult to determine with a 1 decade limit, and also to accommodate children who were able to count to the maximum number possible in the Highest Count task (100). Note that in this work, we applied this criterion to children who made errors in the count list that were not corrected (i.e., mid-decade errors, or errors at decades that were missed by the experimenter), but were still able to recover from those errors. We acknowledge that other reasonable criteria could surely be adopted, too, and that this work is in large part exploratory. For this reason we will share our data on the OSF for others to analyze and subject to alternative criteria. Critically, however, our criteria were specific prior to data analysis, and based on our best guess regarding what a productive rule should predict.

In addition, we have added a section in the Supplemental Information (Section D) where all analyses were repeated using a more conservative criterion that allows for only one error. This resulted in 6 participants re-classified as Non-Productive counters (5% of our total sample). These analyses led to the same conclusions as reported in the main text regarding how Productivity influences Highest Count performance, Next Number performance, and Successor Knowledge. However, there was a different result for regression analyses predicting Endless knowledge and Full knowledge of infinity. First, the original findings reported that Productivity group significantly predicted Endless knowledge; this effect did not replicate using the stricter definition. Second, the original findings did not yield any significant predictors of Full infinity knowledge when controlling for age; however, the stricter Productivity definition significantly predicts Full infinity knowledge, even when controlling for age or Initial Highest Count. Given that Full infinity knowledge requires Endless knowledge, these results generally converge.

In summary, these analyses indicate three robust findings: (1) there exist significant individual differences in knowledge of the decade+unit rule, which affect performance on counting and Next Number tasks; (2) knowledge of this productive rule does not predict successor function knowledge (i.e. the belief that we can always add one); (3) Productivity does predict infinity

[1] 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, 101263. https://doi.org/10.1016/j.cogpsych.2019.101263

understanding (i.e. the belief that numbers never end), although particular regression results may depend on how strictly we define productivity and infinity knowledge (i.e. Endless knowledge vs. Endless+Successor knowledge)

Figure 2 is quite confusing as well. The 6th-10th children from the left in Fig 2a have long orange lines, indicating that they made errors across eight to nine numbers. Why are they still considered productive, if they made more than 3 errors? I understand that it's hard to quantify these rather open ended responses. However, showing the robustness of the results against some of these decision parameters (and the replicability/reliability of these categorization procedure*) must be considered.

*** This should be beyond the scope of this paper, though.**

Thank you for the comment; we have updated the figure (now labeled Fig. 3) for clearer representation of errors. The orange lines represent continuous skipped sequences, which are coded as a single error. The specific participants you mention (participants 46 to 50 in Fig 3a) had skipped the continuous thirty-X sequence, counting: “28, 29, 40, 41, …”. To better indicate that these are a single error, we now use grey lines to represent skipped sequences. Thus, orange dots indicate either individual errors (e.g. “12, 13, 18, 15, 16”) or the beginning of a missed sequence.

In this way, the number of orange dots represent the total number of errors made. Participant 47 is categorized as a Productive Counter according to the improvements they made after receiving Decade Prompts: they counted past their Decade-Change Error at 49 by at least 2 decades after being prompted with ‘50’ and ‘60’.

Our shared dataset includes children’s raw counting sequences with any experimenter prompts. We hope this allows others to explore alternative criteria for classifying children’s responses.

The term "syntax" in the title is misleading because the measure of productivity is rather related to sequence. In fact, the authors appropriately explain it in terms of the "verbal count sequence" in the abstract, and that phrase is more appropriate in the title as well.

We understand why this use of the word syntax might be confusing, and so we’ve now made edits to make our usage clearer. Central to this, our paper is interested in how the words within the counting sequence are generated, rather than how children learn relations between words within the sequence. Our hypothesis is that count words are generated by a syntactic (or morphological) rule. We now take greater care to explain this rule in the paper, and how it relates to infinity. Note that the way we use the words morphology and syntax to refer to counting structures is compatible with how it is used elsewhere in the literature (e.g., Rule et al., 2015; Yang, 2017; Comrie, 2011).

There are issues with the results from the infinity questionnaire. First, the results do not replicate their earlier study: "Prior research suggests that how high children can count is related to their ability to identify successor relations for

known numbers" (p.33). Unfortunately, IHC did not predict Successor knowledge, Endless knowledge, or both combined (Tables 2-4) in this study. These results question the reliability of the test. In fact, the authors themselves acknowledge that children's responses to these questions are extremely unreliable (p.36). Second, it's hard to understand how some children (24%) are only successor knowers and other children (9%) are only endless knowers. Logically speaking, successor knowledge should lead into the concept of endlessness, which leads me to think that 24% of the children ended up being categorized that way because the questionnaire does not truly tap into their understanding of successor or endlessness, or at best provides a very unreliable measure of such knowledge.

Thank you for the comment and for pointing out the two sources of potential confusion. It is important here to clarify what is meant by "successor knowledge" in our study and what is measured in other studies. In most other studies (e.g. Cheung et al, 2017 and Sarnecka & Carey, 2008), knowledge of successor relations is measured by the Unit Task, which asks children to identify number words that represent a quantity one larger than a target set. Thus, the "ability to identify successor relations for known numbers" (p.36) involves knowing that adding one object to a set labeled "four" now elicits the label "five". In contrast, in our study Successor Knowledge on the Infinity Interview (Tables 3-5) refers to endorsing the claim "you can *always* add one to a number". For improved clarity, we updated our manuscript to refer to this as "successor function knowledge", to contrast with item-based knowledge about particular successors.

Due to the difference in defining "successor knowledge" between our study and Cheung et al (2017), we believe that our results are not a failure to replicate. Instead, our results are compatible with Cheung et al (2017) for two reasons. First, as described above, "Successor Knowledge of Infinity" refers to knowledge that "you can always add one", i.e. successors exist for all numbers, without requiring children to verbally identify them. Thus, we think that it is a different construct than the knowledge of specific successor relations tested in Cheung et al on the Unit Task. In fact, results from Cheung et al (2017) support this distinction: "children's performance on the Infinity Task remains relatively poor until they are near ceiling on the [Unit] Task." (Cheung et al, 2017, p.29). Second, while Cheung et al (2017) find that children with full Infinity Knowledge (i.e., both Endless & Successor beliefs) tend to have higher Initial Higher Counts, they did not explicitly test the relationship between IHC and performance on the Infinity Interview while controlling for age. We also find that children with full infinity knowledge have higher IHC (M=68) than other participants (M=46), however the relationship between IHC and Infinity knowledge is not significant after controlling for age (Table 5).

Thus, we believe that our findings do not reflect a failure to replicate and instead reveal additional detail about the relationship between counting and infinity understanding.

The second issue refers to the existence of children who we can label "Successor-only knowers of Infinity" and "Endless-only knowers of infinity". You are right to point out that the claim "you can always add one to a number" (i.e. Successor Knowledge of Infinity) logically implies the claim "numbers never end / there is no biggest number" (i.e. Endless Knowledge of Infinity). In fact, in the paper we acknowledge this logical implication in our discussion of the Peano-Dedekind axioms (p. 11). While this might reflect unreliability in the Infinity Interview, another interpretation would be that 4- to 5-year-old children have not fully recognized the logical implications of the "Successor Knowledge of Infinity" for a fully mathematical understanding of the natural numbers. Very generally, much of mathematics is an exploration of the entailments of first principles, little of which is known automatically, but must be discovered and proved by mathematicians. Support for the idea that this task is likely reliable, and that children learn the entailments of the successor function gradually comes from the compatibility between our findings and previous studies using similar Infinity interviews, including Cheung et al (2017) and Hartnett & Gelman (1998). For instance, Cheung et al (2017) found 33 "Successor-only knowers of Infinity" among a sample of 100 4-7-year-old children. Similarly, Hartnett & Gelman (1998) classified about 50% of children aged 5-6.5 years as "Waverers", which correspond to children with either "Successor-only" or "Endless-only" knowledge of infinity but not both. Finally, both these papers find that older participants were often reliably classified as Full Infinity Knowers, suggesting that variability in Infinity Knowledge status comes from variability in participants' knowledge of infinity concepts.

We have now added text to the paper that clarifies how our work relates to previous reports (to address questions related to replication), and have also noted that although a recursive successor function entails infinity, it is reasonable to expect that children do not automatically compute the entailments of their beliefs.

There is a high correlation between IHC and Productivity Group, even if those with IHC=99 are removed. For example, the majority of the Productive Counters have IHC>=40; the majority of the Non-Productive Counters have IHC<40. Therefore, in their regression analyses (e.g., Tables 2-4), IHC must be entered together with Productivity Group (and the interaction should be assessed as well). Otherwise, it is not possible to tell whether the effect of Productivity Group is due to productivity only or due to productivity combined with IHC.

Thank you for the suggestion. We are also interested and have updated the results section (Section 3.3) to include this logic. Because the only models that yielded a significant effect of Productivity Group was in predicting Endless knowledge, we have updated Table 4 with additional models controlling for IHC and the Productivity:IHC interaction. These additional regression models yielded the same conclusion; Productivity was a significant predictor of Endless knowledge even when controlling for IHC.

We have also included a Productivity by IHC interaction in predicting Next Number accuracy (Section 3.1.2). These results are now updated.

Minor issues:

On page 11, it is said "[children may] infer that numbers must be infinite, even without yet understanding how this recursive rule relates to cardinality" when raising an alternative hypothesis to the idea that successor knowledge must mediate the relationship between counting knowledge and infinity knowledge. That statement is a possibility, but none of the tasks tests cardinal knowledge, so that sentence seems out of place and needs to be revised if not removed.

Thank you for the suggestion; we have now replaced the word “cardinality” with “the possibility of adding 1” (p. 12)

Why are 40 and 70 categorized as a mid-decade number? The step from 40 to 41 is vastly different from the same steps in between 41 and 49.

Our hypothesis was that lacking knowledge of decade words (“forty”, “fifty”) might mask children’s latent knowledge of the decade+unit rule for composing two-digit numbers. Thus, we classified numbers as either cross-decade (child has to generate the correct decade word, e.g. 29 >> 30) or mid-decade (correct decade word is already provided). Since the step from 40 to 41 does not require the child to generate the decade word “forty”, this trial counts as a mid-decade number.

Reviewer #3:

The authors report a study with 4 and 5 year olds, investigating their knowledge of counting (through 3 tasks) and two aspects of "infinity" (successor and endless numbers). They investigate the relations among these measures to shed light on which skills may be more/less associated or accessing distinct aspects of the concepts. The pattern of results in terms of children's performance on the three different counting tasks is interesting in and of itself, but the primary finding (or, what I take as the author's primary finding) is that they do find evidence of distinct relations between productive counting and successor vs. never-ending numbers — suggesting that these two aspects of infinity are differentially related to the counting structure knowledge.

The inclusion of these tasks in the same children provides important theoretical and methodological contribution to the study of children's number knowledge. However, the are aspects of the manuscript that are difficult to follow or otherwise lower the potential impact. My more detailed comments are below:

Method

In general the method was clear, although the coding decision-making sections were sometimes difficult to follow. I wonder whether a decision tree (even just in supplemental) would be helpful to more clearly communicate how these bins were determined.

Thank you for the suggestion; we have added a decision tree (Fig. 1) to the manuscript to describe the Highest Count coding scheme.

Also, can you provide more information/justification for the 3 errors cut off? Why 3 or fewer vs. 4 or more errors? What was the distribution of error numbers? (i.e., did it tend to be 1 or 2 vs. a lot? or were there many children at those border cases of 3 vs. 4 errors, which seems rather arbitrary?)

Thank you for the comment. We have now added details to the paper, and also describe our logic above in response to Reviewer 2's question above, which is based on the logic of previous, pre-registered analyses.

- Also, of the children that made lots of errors, did they tend to be the same error (e.g., always says 5, 4, 6 — 25, 24, 26 — 35, 34, 36 etc.) — because if so, this would result in 9 errors, but seems unfair to call them not productive counters.

- More descriptive information about these categories, the number of / distribution of children within the categories for various reasons (i.e., because the made no errors vs. 1 error etc), would help clarify and provide some justification for why 3 errors was used as the cut-off

- Actually, now having read further, Figure 2b can speak to this already a little bit, so it may be helpful to refer to this figure to discuss whether there were any/many cases that were given an "arbitrary" category, e.g., 4 errors vs. 3 errors. Numbering the participants to refer to specific borderline cases or which cases went into the given category for which reasons could be helpful for understanding the justification of this cut off.

Thank you for the suggestion; We have added participant indexes to Fig 2 (now renamed Fig 3).

In regards to the distribution of children's errors, we did not find any children who made the same error that repeated across decades. Children who made many errors tended to recite numbers in random sequence. Participants' raw count sequences can be found in our published dataset, available on the OSF repository.

In addition, most participants made 3 or fewer errors; only one participant made 4 errors. For participants who counted to 100 on their own, only one participant made 4 errors (Fig 3b., subject 23). Participants who were classified as unproductive either stopped counting on their own in the middle of a decade, or stopped counting after failing to continue counting past the first or second decade prompt (n=18).

Results:

In general, the results are difficult to follow, but the figures and tables greatly help the reader. It would be helpful to future subset the results beyond just the task being analyzed (especially since section 3.1.2 Number Number Task, for example, also includes substantial analysis involving the counting tasks as well). For example, organized by research question or categories like "Describing Performance" vs. "Relations Among Tasks" would help the reader navigate the substantial information being provided.

We have renamed section headings: "3.1.2: Productivity and Next Number performance"

Page 19: "However, many Productive Counters (n=9) had an Initial Highest Count of 29 or lower" It's not clear that 9/73 productive counters would be considered "many". It's an important note that there are some Productive Counters with low IHCs, but Given the subjective interpretation of "many", it may be better to simply say "some"

Thank you for the comment; we have edited that sentence for clarity.

- Section 3.1.2: first analysis is a t-test with all children, second is a mixed effects regression with a subset of children —> it's a little hard to follow why the change in analytic approach. The explanation for why looking at a subset of children is clearly provided, but the motivation for changing the analytic strategy is unclear. I think this might just be an issue of more fully explaining the rationale, but it may also need some alignment of analytic approaches

The general analytic approach in this paper is the use of generalized linear (mixed effects) models. We have edited the text in Section 3.1.2 to explain that the first t-test analysis provides only a descriptive comparison of productive and non-productive counters without controlling for covariates of interest (e.g. age and initial highest count).

- The paragraph of discussion of this section on Page 23 is also very hard to follow, it would be helpful to provide a bit more information about what analyses or aspects of the methods these claims are following from (e.g., the idea that productive counters could have only been tested on small numbers was jarring, until I realized that was because they could have erred on 20, and then gone as far as 40 without another error; but it would be helpful to spell that out more in this section; the "first analysis" isn't clear - do you mean the t-test comparing Productive and Not Productive? If so, it's unclear how that suggests that Next Number and Productivity capture different aspects of counting knowledge.

Thank you for the comment; we have edited the text so that Section 3.1.2 clearly describes 3 separate analyses (t-test, regression controlling for IHC and age, regression including specific items). We have also edited the discussion paragraph (page 37) to refer to these specific analyses and explain which claims follow from which analyses.

- it would be useful to point to tables 2-4 earlier in the results section, as it's much easier to follow the narrative explanation of the results along with the tables; for e.g., indicating that results are in Table 2 when stating: "Interestingly, in predicting children's Successor Knowledge, none of the three predictors explained a significant proportion of additional variance compared to the base model."

Thank you for the suggestion; we have referred to these tables earlier in the results section (p.32). We also refer to the relevant table when describing each set of regression analyses by outcome variable (Successor, Endless, or Full Infinity knowledge).

Other Major Comment:

I highly recommend the authors make their data, analysis code, and/or materials available on a repository (e.g., the open science framework, although there are also others) so that others can easily replicate and reproduce their results. Notably, this would be particularly useful given my question about the categorization of individuals into productivity bins - if others are unsure of this categorization, they can more closely inspect or re-analyses with different decisions, allowing others to investigate the robustness of these results in other ways

We have made our data, analysis code, and experimental protocol available on the Open Science Framework: https://osf.io/z6ky3/

Counting to Infinity: Does learning the syntax of the count list predict knowledge that numbers are infinite?

Junyi Chu[a,1,*], Pierina Cheung[b], Rose M. Schneider[a], Jessica Sullivan[c], & David Barner[a]

[a]University of California San Diego, La Jolla, CA 92093, USA

[b]National Institute of Education, Nanyang Technological University, Singapore

[c]Skidmore College, Saratoga Springs, NY 12866, USA

Word Count: 10797

* Corresponding author.

Email address: junyichu@mit.edu

[1] Present address: Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02138, USA

Abstract

By around the age of 5½, many children in the US judge that numbers never end, and that it is always possible to add +1 to a set. These same children also generally perform well when asked to label the quantity of a set after 1 object is added (e.g., judging that a set labeled "five" should now be "six"). These findings suggest that children have implicit knowledge of the "successor function": every natural number, *n*, has a successor, *n*+1. Here, we explored how children discover this recursive function, and whether it might be related to discovering productive morphological rules that govern language-specific counting routines (e.g., the rules in English that represent base 10 structure). We tested 4- and 5-year-old children's knowledge of counting with three tasks, which we then related to (1) children's belief that 1 can always be added to any number (the successor function), and (2) their belief that numbers never end (infinity). Children who exhibited knowledge of a productive counting rule were significantly more likely to believe that numbers are infinite (i.e., there is no largest number), though such counting knowledge wasn't directly linked to knowledge of the successor function, *per se*. Also, our findings suggest that children as young as four years of age are able to implement rules defined over their verbal count list to generate number words beyond their spontaneous counting range, an insight which may support reasoning over their acquired verbal count sequence to infer that numbers never end.

Keywords: Count list, Infinity, Conceptual change, Successor function, Highest count, Decade+Unit rule

1. Introduction

Human learners draw on a finite set of experiences to acquire information about the world, but nevertheless acquire systems of rules that permit the generation of unbounded representational outputs. For example, natural language is often touted as an example of how humans make “infinite use of finite means”: a finite lexicon and system of combinatorial rules allows children to generate an unbounded number of possible utterances (Chomsky, 1965; Humboldt, 1836/1999, p.91). Similarly, numerate humans learn a set of symbols and combinatorial rules that permit an unbounded set of mathematical expressions. For example, the English base-10 numeral system expresses 80 numbers (from twenty to ninety-nine) by composing just 17 unique decade and unit words (twenty through ninety, and one through nine). The expressive power of this system is unbounded: to express larger numbers, one simply needs to learn the appropriate words representing powers of ten (e.g., hundred, thousand, million, etc.), and recursively compose them following the syntactic rules of numerals (Cheung, Dale, & Le Corre, 2016; Hurford, 1975). As we describe below, although children initially believe that numbers are finite – and have limited knowledge of both the structure and meaning of number words – they ultimately come to believe that numbers never end. In the present study, we investigate how children learn that numbers are infinite, and whether the rule-governed structure of the verbal count list might play a role in children’s inference that numbers form an infinite class.

When children first begin learning about numbers in early childhood, their knowledge is clearly item-based and finite. At around the age of two, English-speaking children in the US begin to recite a subset of the verbal count list (one, two, three, four, etc.), but often can't count beyond ten at this point (Fuson, 1988; Fuson & Hall, 1983; Gelman & Gallistel, 1978).

Moreover, these number words appear to lack meanings at this early stage: When asked to give a number (e.g., to give one fish), children initially give a random amount (e.g. Schaeffer, Eggleston, & Scott, 1974; Wynn, 1990). Some months later, children appear to acquire an exact meaning for the word one, and can give one object when asked, while failing to reliably give two when asked. These children are often called one-knowers. Another 6 to 9 months later the children become two-knowers (and can reliably give two objects), then three-knowers several months after that. One by one, children add meanings to their number words in a way that suggests the lack of a productive logic governing these meanings (Sarnecka & Lee, 2009).

While children initially lack a productive rule for understanding number words, a breakthrough appears to happen at around the age of 3 and a half or 4 (in US English-speaking groups), when children appear to realize that they can correctly give any requested number by counting and giving all objects that are implicated in their count. These children are often called "Cardinal Principle Knowers" or CP-knowers. While the time-course varies across language groups, the basic developmental sequence - i.e., of progressing through discrete number-knower stages and finally using the counting routine to productively generate any quantity within the child's count list - has been reported for children learning a variety of languages around the world (Almoammer et al., 2013; Barner, Libenson, Cheung, & Takasaki, 2009; Le Corre, Li, Huang, Jia, & Carey, 2016; Piantadosi, Jara-Ettinger, & Gibson, 2014; Sarnecka, Kamenskaya, Yamana, Ogura, & Yudovina, 2007). Also, in bilingual preschoolers, children progress through discrete number-knower stages independently in each language, but generally become CP-knowers at the same time in both languages (Wagner, Kimura, Cheung, & Barner, 2015). What remains unclear, however, is what, exactly, children learn when they become CP-knowers.

By some accounts, becoming a CP-knower not only involves mastery of an enumeration procedure for set sizes within one's familiar count list, but also an inductive leap in understanding the meaning and structure of natural numbers. Multiple researchers have argued that, to learn how counting works, children construct a type of analogical mapping between the verbal count list and the ordered set of cardinalities that the list represents, beginning with the numbers one, two, and three (Carey, 2004; Gentner, 2010; Schaeffer et al., 1974; Wynn, 1992; for review, see Marchand & Barner, 2018). For example, Wynn (1992) argues that "in order to learn the counting system, children must implicitly make the analogy between the magnitudinal relationships of their own representations of numerosities, and the positional relationships of the number words." (p. 250). Similarly, according to Carey (2004): "Children may here make a wild analogy– that between the order of a particular quantity within an ordered list, and that between this quantity's order in a series of sets related by additional individuals. If the child recognizes this analogy, they are in a position to make the crucial induction: For any word on the list whose quantificational meaning is known, the next word on the list refers to a set with another individual added." (p. 67). Following this logic, Sarnecka and Carey (2008) created a measure that they called the Unit Task, in which children were told, e.g., "Ok. I'm putting FOUR frogs in the box", then saw 1 or 2 items added, and were then asked, "Now is it FIVE or SIX?" (trials included 4+1, 4+2, 5+1, 5+2). They found that children identified as CP-knowers by Wynn's Give-a-Number task succeeded on 67% of trials overall, whereas subset knowers performed at chance. Based on this, they concluded that becoming a CP-knower involves more than acquiring a procedural rule, and instead marks the moment at which children acquire the successor function.

On this account, the ability to accurately count objects therefore reflects the ability to link cardinal representations with implicit knowledge of a type of logical rule, called the "successor function", described by logicians and philosophers of mathematics including von Leibniz (1704/1996), Peirce (1881), Dedekind (1888/1963), and Peano (1889) in efforts to construct axioms to define the natural numbers. One such system, commonly referred to as the Peano-Dedekind axioms, include principles akin to those in (1), inter alia (though various notational variants exist):

(1) i. 1 is a natural number.

ii. All natural numbers exhibit logical equality (e.g., x=x; if x=y, then y=x, etc.).

iii. For every natural number n, S(n) (the successor of n) is a natural number.

iv. Every natural number has a successor.

Critically, because the Peano axioms state that every natural number has a successor, they generate an infinite number of numbers. Consequently, a child who has implicit knowledge of such rules would be expected to believe that it is always possible to add 1 to a number, and also that numbers never end. Thus, this account predicts that becoming a CP-knower represents a shift from representing numbers as a finite sequence of individual words to understanding them as products of a rule - the successor function - that generates an infinite set of positive integers (for discussion, see Sarnecka & Carey, 2008).

Does becoming a CP-knower involve learning to reason about cardinalities in terms of a recursive successor function? The results of Sarnecka and Carey (2008) leave open this question, since they don't test whether children who succeed on the Unit task generalize this knowledge to all numbers in their count list, let alone to all possible numbers. Also, while they showed that CP-knowers outperformed subset knowers on the Unit task, they didn't show that success on the

Unit task, and thus the ability to reason about cardinalities in terms of successor relations, was a prerequisite to becoming a CP-knower. Instead, they tested only very small numbers - four and five - leaving open the possibility that children's knowledge was item-based, rather than governed by an abstract logical rule. Also, the study left open the precise timeline according to which successor function knowledge is acquired.

Several subsequent studies have suggested that, if children acquire a rule akin to the successor function, this likely occurs several years later than predicted by Sarnecka and Carey (2008). First, Davidson, Eng, and Barner (2012) tested a large group of CP-knowers with the Unit Task, but included a slightly wider range of numbers, extending from 4 to 25. As a proxy for experience with number words, Davidson et al. asked children to count as high as they could, and binned them into low, medium, and high counters, analyzing Unit Task performance only for numbers within each child's count list. They found that almost all low counters (who could count up to 19) performed at chance on the Unit Task for numbers within their count range despite being CP-knowers, and that only the highest counters (who could count beyond 30) performed systematically well on small numbers. For larger numbers, all groups performed relatively poorly, even when those large numbers were well within their counting range. Similar results were found in a study of bilingual learners (Wagner et al., 2015), and in a training study which found that many CP-knowers lacked successor function knowledge (though CP-knowers were more likely than subset knowers to show improvement over 2-3 weeks of training; Spaepen, Gunderson, Gibson, Goldin-Meadow, & Levine, 2018). Also, children perform still poorer when asked to reason about predecessors in the Unit task (though this may reflect the working memory challenge associated with counting backwards, rather than children's understanding of how moving up and down the count list relates to number; see Kaminski, 2015; Sella & Lucangeli, 2020). Finally, in a study of slightly older children, Cheung, Rubenson, and Barner (2017) found that CP-knowers did not perform reliably above chance on the Unit Task for all numbers in their count list until around the age of 5 and a half - almost two years after most children become CP-knowers. In sum, these studies suggest that children learn to reason productively about successor relations only after a protracted period of item-based learning. This raises the possibility that learning about successor relations depends on some gradual inductive inference or additional knowledge about count words that are not captured in these studies.

Critically, as noted above, the successor function doesn't merely state that for a particular number, *n*, there is a successor. Instead, it states that every number has a successor, such that numbers are infinite (for similar treatments of the successor function, see e.g. Decock, 2008; Wright, 1983). Given that knowledge of particular numbers could plausibly reflect memorized knowledge rather than the application of a productive successor function, Cheung et al. (2017) paired the Unit Task (which tested how children implement the successor function in particular numbers) with an infinity interview first reported by Gelman and colleagues (e.g., Evans, 1983; Hartnett & Gelman, 1998), which tested children's beliefs about numbers as a class. In this battery, children were asked about the largest number they could name and whether it was the largest possible number, or whether it might be possible to repeatedly add 1 to it. This "successor question" tested whether children believe that numbers can be generated via a +1 rule. Children also answered an "endless question" about whether counting would get them to the end of numbers, or if numbers went on forever. Like earlier studies on this topic (Evans, 1983; Hartnett & Gelman, 1998), Cheung et al. (2017) found that children initially believe that numbers are finite, and that it's not always possible to add 1, but that by around the age of 6 many undergo a transition and begin to claim that numbers never end. Further, they found that this knowledge emerged shortly after children were able to reason about successor relations and pass the Unit task for large numbers in their count list.

How might children learn, from knowledge of a finite count list, that numbers exhibit a successor function, and are infinite? Previous studies have found that children's ability to identify successors for known numbers, as measured by the Unit task, is related to how high they can count (Cheung et al., 2017; Davidson et al., 2012). For example, Cheung et al. (2017) found that children who could count up to at least 80 (many of whom could count even higher) were able to identify the cardinal value of successors for a wide range of known numbers within their count list, whereas children with lower highest counts could only do so reliably for the smallest numbers. This observation suggests at least two mutually compatible explanations for the relationship between counting experience and successor function knowledge. The first possibility is that there is no direct link between how high a child can count and successor knowledge, and that these two outcomes are correlated because they both result from general exposure to number. The second possibility is that there is a more direct causal link between the two: that children's understanding of how count words are syntactically structured might inform their intuitions regarding successor relations and infinity, and that this structure is only apparent after children have learned to count to relatively large numbers. Specifically, Cheung et al. (2017) noted that when children learn to count in English, they are required to learn a recursive base 10 structure wherein they first count from one to nine, then repeat this one through nine structure with varying degrees of regularity for higher decades, which themselves are generated by multiplying 1-9 by 10 (for related proposals, see Barner, 2017; Hurford, 1987; Rule, Dechter, & Tenenbaum, 2015; Yang, 2016). Compatible with other cases of morphological learning (e.g., the past tense or plural; Pinker & Ullman, 2002), children may begin by simply memorizing items in their count list, but transition to a system of rules when the number of regular items is sufficiently large, resulting in a rule that can generate an unbounded number of exemplars (and in tandem, the intuition that numbers may never end). For example, after having learned to count to forty or fifty, a child might notice that after each decade term (twenty, thirty, forty, etc.) the next number can be generated using an additive decade+unit rule, by adding the words one through nine in order as in Table 1:

Table 1.
Example of Decade+Unit Rule for "thirty" and "xty" and corresponding cardinalities

Decade label	Unit label	Decade+Unit Label	Cardinal value (defined by successor function)
thirty	one	thirty-one	31 = 30+1
	two	thirty-two	32 = 31+1
	three	thirty-three	33 = 32+1
	four	thirty-four	34 = 33+1
	n	thirty-n	3n = 3m+1
	…	…	…
xty	n	xty-n	xtyn = xtym+1

Consistent with this hypothesis, previous studies have found that children who can count beyond 100 are better able to decompose numbers into decades and ones, whereas children who can't count as high appear to store the count list as a memorized string (Fuson, Richards, & Briars, 1982; Siegler & Robinson, 1982). Also relevant is that children who can't yet count all the way to 100 nevertheless make errors which suggest some knowledge of rules that structure counting. For example, when asked to count as high as they can, many children stop at decade transitions (Fuson et al., 1982; Siegler & Robinson, 1982; Wright, 1994), with the most frequent being 29 and 39 (Gould, 2017). If children were merely reciting a memorized and unstructured list as they do the alphabet, we might expect the distribution of their errors to be random rather than at decade transitions. Instead, their errors suggest that children have memorized an initial list, e.g., up to 20 or 30, and use some form of morphological rule, like the one described above, to generate numbers up to the next decade transition (which requires memorized knowledge, since in English decade labels are irregular, and can't be generated from a rule alone). Consistent with this, children exposed to languages with relatively transparent base-10 counting systems, like Mandarin or Cantonese Chinese, appear to count higher and make fewer errors than children learning less transparent counting systems, like English or Welsh (Miller, Smith, Zhu, & Zhang, 1995; Miller & Stigler, 1987; for related work, see Dowker, Bala, & Lloyd, 2008). Such evidence suggests that children make use of the linguistic structure of their count list to learn rules governing counting. An open question - and the main focus of the present study - is whether learning that number words are compositionally structured might facilitate insights into the conceptual structure of numbers, such as learning the successor function and infinity.

Critically, although the successor function logically entails that numbers are infinite, young children may not automatically compute the entailments of their beliefs, and may not infer from their successor knowledge that numbers never end. As noted by Cheung et al. (2017), many children in their study believed that it's always possible to add 1 to a set but nevertheless believed that numbers must ultimately end (children they called Successor Only Knowers), whereas only a handful of children held the opposite pattern of beliefs - that you can't always add 1, but that numbers are nevertheless infinite (what they called Endless Only Knowers). Previous studies find the same pattern but report no Endless Only knowers at all (Evans, 1983; Hartnett & Gelman, 1998). Cheung et al. (2017) interpreted this pattern as evidence for a developmental sequence whereby children learn some kind of bounded (item-based) successor rule that applies to a finite list, and only later learn that numbers never end. For example, children may first learn that known numbers exhibit a successor relation by empirically noticing this relation between familiar numbers, but may make the induction that this function is recursive by learning that *number words* can be productively generated via the additive decade+unit rule that governs counting. On this hypothesis, experience with counting may be related to both learning the successor function and that numbers are infinite, but for different reasons – e.g., in the first instance providing opportunities for item-based observation regarding relations between known numbers, and in the second instance allowing children to discover rules that generate an unbounded set of numerical symbols.

While previous studies have shown that how high a child can count is related to their ability to identify successor relations of specific numbers (Cheung et al., 2017; see also Davidson et al., 2012), they have not tested whether knowledge of productive counting rules is related to the belief that numbers are endless, independently of the belief that every number has a successor. This is important, because learning rules that govern the structure of number words may allow children to make a broad inference about numbers – e.g., by learning that number words can be decomposed into decades and ones, children may realize that such rules can generate an infinite set of number words, which may in turn form the basis for the belief that numbers are infinite. Notably, although this relation between counting and infinity might hinge on first acquiring knowledge of successor relations, this needn't necessarily be the case. For example, children might learn that it's possible to generate indefinitely many numbers based purely on the syntax of the count list, and thereby infer that numbers must be infinite, even without yet understanding how this recursive rule relates to the possibility of adding 1. Given this possibility, it's important to test how children's knowledge of the structure of counting relates both to learning that it's always possible to add 1 to a number, and separately to learning that numbers are infinite.

In the present study, we had two goals. First, we sought to measure children's acquisition of the "decade+unit" rule to determine when such knowledge emerges, and how it is related to general counting experience (e.g., highest count). Second, we asked whether knowledge of such a rule was related to both acquisition of successor function knowledge and the belief that numbers are infinite. To do this, we presented 4- and 5-year-old children with three tasks.

First, to assess children's counting productivity, we tested them with the "Highest Count" task. A critical difference between this study and previous work is how we evaluated children's highest count. Specifically, in previous studies of successor function knowledge, a child's highest count was defined as the highest number to which they could count before their first error. While this is likely a good first-pass measure of how much training children have received, it leaves open whether a particular child has acquired a productive rule or has simply memorized their count sequence. For this reason, we explored not only how high children could count without error, but also which number they stopped on, and what they did when provided with the next number. As already noted, previous studies find that children's highest counts are not randomly distributed, and that instead their most common first error occurs at decade transitions, compatible with having learned a decade+unit rule. To test whether children count up to decades (e.g., up to 29, 39, or 49) by exploiting such a rule, we provided children who stopped at a decade transition with the next decade term (e.g., 30, 40, or 50) and asked whether they could then continue counting. We reasoned that children who have acquired a productive decade+unit rule should be able to count higher once decade terms are provided - i.e., they should know that for any decade label N-ty, the next number in the counting sequence should be N-ty-one, followed by N-ty-two, etc.

Second, as an alternative test of whether children have acquired a productive rule for generating numbers, we used the Next Number task, in which children were told a number (e.g., “fifty-seven”) and asked to generate the next number in the count sequence (i.e., “What comes next?”). We reasoned that children who understand the decade structure of counting should not merely represent the count list as a single memorized string, but should be able to generate the next number for any decade - i.e., they should exhibit knowledge of how the verbal count list implements the successor function. Based on this, we reasoned that productive knowledge of counting as measured by the Highest Count task should be correlated with performance on the Next Number task.

Finally, to explore the second goal of this study, we asked whether either of these two tests of counting productivity predicted children’s intuitions about infinity in a qualitative Infinity Interview. It is important to ask children questions that are not tied to specific numbers, because a child might be able to label the next number for a finite list of words, but not know the successor function - i.e., that *every* number has a successor. In particular, we tested how our counting measures were related to children’s belief that it’s always possible to add 1 (i.e., successor function knowledge), and that numbers never end. Multiple past studies have reported ceiling performance on this task by age 6-7 years (Cheung et al, 2017; Evans, 1983; Hartnett & Gelman, 1998); thus, by studying younger children (ages 4 and 5) we investigated what factors contribute to variability in the acquisition of infinity knowledge.

2. Method

2.1 Participants

We tested 122 4- and 5-year-old children (M = 5;0, SD = 7 months, range = 4;0 to 5;11, 58 male) recruited from preschools and museums in the San Diego metropolitan area. Although we did not collect demographic information from each participant, our sample was drawn from a population with the following statistics: White (75.5%), Black (5.5%), Asian (12.6%), American Indian or Alaska Native (1.3%), Pacific Islander (0.6%), Multiracial (4.5)%.[2] All participants spoke English as a primary language. A stopping rule was defined such that 30 children were tested in each 6-month age bin within this range of ages. An additional 25 children participated but were excluded for the following reasons: their primary language was not English (*n* = 4), they did not complete the tasks (*n* = 10), parental interference (*n* = 1), experimenter error (*n* = 3), or failing the practice trials on the Next Number Task (*n* = 9).

2.2 Stimuli and Procedure

2.2.1 Highest Count Task. We used the Highest Count task to measure children's spontaneous counting ability, and to classify children as either Productive or Non-Productive Counters based on their ability to recover from errors on decade transitions. Below, we describe procedures in detail. Fig. 1 shows a simplified decision tree describing the protocol and classification rules for determining productive counting knowledge.

To test children's knowledge of the counting sequence up to 99, we first asked each child to count as high as they could. If the child failed to respond, the experimenter said, "Let's count together! One..." with rising intonation to encourage the child to continue counting alone. We allowed children to count until they stopped naturally, and recorded errors made along the way.

[2] For detailed demographic information regarding San Diego residents, see https://www.census.gov/quickfacts/fact/table/sandiegocountycalifornia,CA/PST045218

Each skipped number (e.g., “12, 13, 15”), skipped sequence of numbers (e.g., “18, 19, 30, 31”), or substitution error (e.g., “4, 9, 6”) was counted as one error. Children were also allowed to self-correct or restart counting with no penalty (e.g., if a child counted “1, 2, 4, no, 1, 2, 3, 4, 5”, 3 would not count as an error). To avoid underestimating children’s counting ability due to lapses in attention, we always reminded children of the last number they had said whenever they stopped counting. For instance, a child who stopped at 25 was prompted with, “So what’s after 25?” Children were allowed to continue counting and receive as many reminders as necessary; these pauses and reminders were not considered errors. Using this method, we obtained children’s Initial Highest Count (IHC), which was the highest number children counted to before making any errors, either on their own or with reminders. We also obtained children’s Final Highest Count (FHC), which was the highest number children ever counted to that was part of a 3-number consecutive sequence. We allowed for up to 10 errors, with at most 3 errors in a single decade. For some children, their Final Highest Count also included experimenter-provided decade prompts, as described below.

In this study, we were especially interested in testing whether children have a productive decade+unit rule for counting, as measured by their ability to either (a) count to 99 on their own with minimal errors, or (b) extend their count sequence when provided with decade labels beyond their initial highest count. To measure the latter ability, during testing we first identified instances where children’s first or second error occurred at a decade transition (e.g., stopping at 29, substituting ‘30’ with ‘twenty-ten’, or skipping 30 altogether). We called these Decade-Change Errors and provided children with a corrective decade prompt. For instance, a child who made an error after 29 was told, “After 29 is 30. Can you keep counting? 29, 30...” Children who

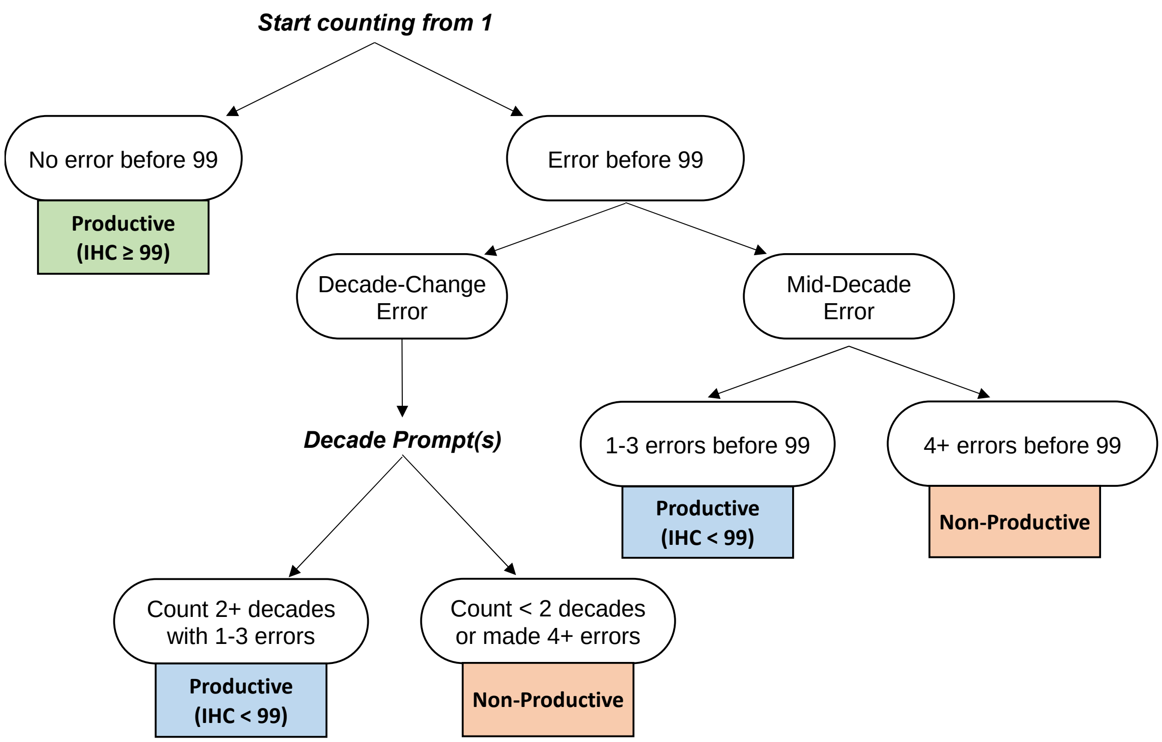

Figure 1. Decision tree for Productivity Classification based on Highest Count Task. Participants started counting from 1 and were classified as either Non-Productive, Productive (IHC<99), or Productive (IHC ≥ 99) counters.

successfully counted-up were provided with decade prompts for any subsequent Decade-Change errors they made, with decade prompts ranging from 20 to 90. Children who made no Decade-Change Errors did not receive any decade prompts and were simply allowed to continue counting until they stopped naturally.[3] The task ended whenever a child reached 99 or if they said they could not continue counting, whichever was earlier.

In addition to obtaining children's Initial and Final Highest Counts and counting errors, we also used the Highest Count task to classify participants as Productive or Non-Productive

[3] Upon coding the data, we noticed that some counting errors were not detected during the experiment, such that some children (n = 10) received a decade prompt despite making mid-decade errors before their first Decade-Change Error. We classified these children as Non-Productive Counters irrespective of their subsequent counting progress, since they should not have received a decade prompt. Any counting past the erroneously provided decade prompt was excluded from our analyses.

Counters. To classify children, we used criteria developed in a previously published, pre-registered report (Schneider et al., 2020). In particular, we reasoned that if a child has a productive decade+unit rule, then they should be able to count up to decade labels, but may not know what those decade labels are, and therefore should be expected to make errors on decade transitions (e.g., 39 to 40). We therefore classified children as Productive if (1) they counted to 99 with three or fewer errors, or (2) they could count at least two decades beyond their initial Decade-Change Error without making more than three errors in those two decades, including any decade-transition errors that elicited a decade prompt. This three-error criterion was developed because it allowed for two decade transition errors, along with a maximum of one mid-decade error. For example, a child whose initial error was at 29 but continued counting with decade prompts to 49 or higher was classified as a Productive Counter, but if they continued to only 39 or made too many errors before getting to 49, they were classified as a Non-Productive Counter.[4] Note that although alternative criteria are possible, we resisted exploring the full space of results that might result from different cut-offs. Still, in response to a reviewer request, in the Supplemental Materials we provide one example of how a more conservative threshold (allowing just 1 error to Productive Counters) affects analyses. We show that this criterion generates a similar pattern of results, with only 5% of children classified differently (see Supplemental Materials D for details).

[4] Because we consider skipped sequences as single errors, this classification allows for the possibility that children improve by two decades with a single error that skips most of that sequence (e.g. “29, [prompt 30], 48,49”). Post-hoc checks confirmed that no participants were classified as Productive Counters due to this loophole; all Productive Counters counted-up correctly from decade prompts (i.e. “[prompt 30], 31”).

2.2.2 Next Number Task. In this task, children were provided with a number and asked, "What comes next?". Children received two practice items (*one*, *five*) to ensure they understood the task, and corrective feedback was provided on these practice items if needed. For example, when asked, "*Five*. What comes next?", children who answered *four* were invited to count and figure out the correct answer (e.g., "No, *four* comes before *five*. What comes after *five*? Can you count and find out?"). All children in the final dataset successfully answered these practice questions before continuing.

Since we were interested in individual differences between children, all children received the same test items in a fixed order, ranging from 20 to 90: 23, 40, 62, 70, 37, 29, 86, 59. These test items were designed to cover the range of counting abilities up to 99 and to utilize the decade rule for forming number words. No feedback was provided during the test trials.

2.2.3 Infinity Interview. Following the protocol of previous studies of infinity knowledge (Cheung et al., 2017; Evans, 1983; Evans & Gelman, 1982; Hartnett & Gelman, 1998), we also assessed children's understanding of infinity by probing two types of belief: (1) that there is no biggest number, and (2) that it is always possible to add 1 to any number. To probe this, we asked six questions as follows:

1. *"What is the biggest number you can think about?"* If the child did not answer, the experimenter probed them by asking how high they could count.
2. *"Is that the biggest number there could ever be?"*
    a. If yes, move on.

   b. If no, "*Can you think of a bigger number? Is that the biggest number there could ever be?"* The experimenter repeated this exchange up to 4 times or until the child affirmed that they had produced the biggest number.
3. "*If I keep counting, will I ever get to the end of numbers, or do numbers go on forever? Why?*"
4. "*If we thought of a really big number, could we always add to it and make it even bigger, or is there a number so big we couldn't add any more? Why / Why not?*"
5. "*You said the biggest number you know is X. Tell me, is it possible to add one to X, or is X the biggest number possible? Why?*" For this question, X was the largest number the child had stated in the entire testing session.
6. "*Could I keep adding one? Why / Why not?*"
   a. If yes, "*What would happen if I kept adding one?*"

Each child was assigned a binary classification for each of two aspects of infinity understanding. Classifications were assigned by the first author and a coder blind to the hypotheses. The coding scheme was consistent with previous studies (Cheung et al., 2017; Evans, 1983; Hartnett & Gelman, 1998) and is provided in the Supplemental Materials (A), along with example transcripts. First, we coded whether children believed that there was a highest number, such that numbers must end, or whether they believed that numbers go on forever, which we labeled as "Endless" knowledge of infinity. This coding was based on responses to questions 1-3 and 5. Second, we coded whether the child believed that it was always possible to add 1 to any number according to their responses to questions 4-6. We call this the "Successor" knowledge of infinity. Initial agreement was 84.0% for Endless Knowledge coding (Cohen's Kappa = .63, $p < .001$), and 80.9% for Successor Knowledge coding (Cohen's Kappa

= .62, p < .001). Disagreements were resolved through consulting a third coder. Finally, we identified children as "Full Infinity knowers" if they endorsed both aspects of infinity understanding.

3. Results

Compatible with the first goal of this study, our first set of analyses examined children's acquisition of the productive decade+unit rule using data from the Highest Count and Next Number tasks. Second, compatible with our second goal, we analyzed children's performance on the Infinity interview, and asked how their responses to the Successor Knowledge and Endless Knowledge items were related to counting experience and productive knowledge of the decade+unit rule.

All analyses were conducted in R (version 3.6.0, R Core Team, 2019). Regression models were constructed using either the R base stats package or, for models containing mixed effects, using lme4 (Bates, Mächler, Bolker, & Walker, 2014). For ease of interpretation, predictor variables were mean-centered for analyses.[5] To test for the significance of specific independent variables, we conducted Likelihood Ratio Tests comparing models with and without particular effects of interest. For analyses containing within-subject measures, we constructed mixed-effects models containing random intercepts for all relevant grouping units (e.g., subject, item).

3.1 Characterizing Decade+Unit Productivity

3.1.1 Productivity and Counting ability. Fig. 2 presents the distribution of children's Initial Highest Count by productivity classification. Overall, 73 children were classified as

[5] Continuous variables were mean-centered and scaled by 1 standard deviation. Categorical variables (all binary in this paper) were also mean-centered and weighted by their group counts (i.e. weighted effect coding, see Grotenhuis et al., 2017). This allows regression coefficients to be interpreted as standardized main effects and to be compared across models.

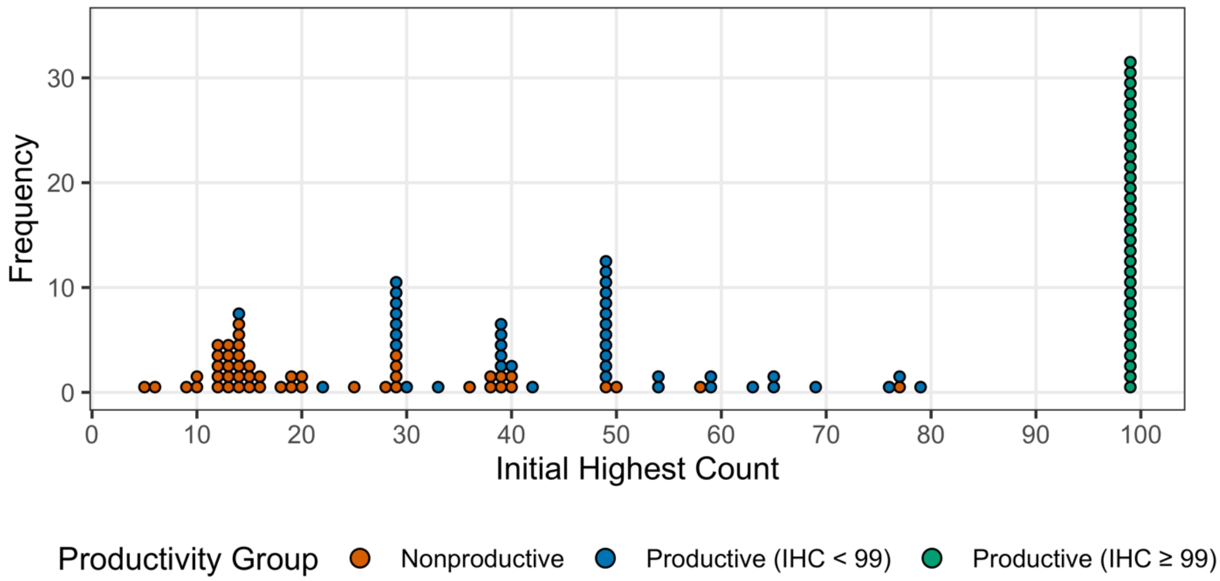

Figure 2. Distribution of children's Initial Highest Count by Decade Productivity.

Productive Counters (M_{age} = 5;3, *SD* = 6 months) and 49 were classified as Non-Productive Counters (M_{age} = 4;7, *SD* = 5 months). Thirty-two children reached 99 on their Initial Highest Count and were therefore classified as Productive Counters (IHC ≥ 99). Of the remaining ninety participants, 49 children (54%) made a Decade-Change error on either their first error (*n* = 37) or second error (*n* = 12) and thus received Decade Prompts. Of these, 32 (65%) continued counting two decades further with fewer than 2 additional errors, and were classified as Productive, while 17 did not, and were classified as Non-Productive. Another 9 children were classified as Productive for counting past their Initial Highest Count to reach 99 with up to three errors, without receiving any Decade Prompts. Note that this pattern of results is similar if a more conservative criterion of just 1 error is adopted. Using this criterion, 6 participants (5% of our sample) are reclassified from Productive to Non-Productive (for details, see Supplemental Materials D).

Productive Counters had, on average, a higher Initial Highest Count (*Mean* = 69.2, *SD* = 28.8, *Median* = 65) than Non-Productive Counters (*Mean* = 22.6, *SD* = 14.9, *Median* = 15). This difference remained if we considered only Productive Counters with Initial Highest Count below 99 (*Mean* = 46.0, *SD* = 15.3, *Median* = 49). However, some Productive Counters (*n*=9) had an Initial Highest Count of 29 or lower.

Next, we analyzed precisely how far children could count past their initial error. Fig. 3 shows participants' Initial Highest Count, Final Highest Count, any errors made, and provided decade prompts, if any. While our classification criteria for productive counting only required children to count past their Decade-Change errors by two decades, most of the Productive Counters were able to count several decades further, and many of them reached 99 (*Mean number of prompts* = 3.5, *SD* = 1.6, *Range* = 1 to 7). Almost half of the Productive Counters (*n*=32) received decade support on the Highest Count Task, and the remaining children (*n*=41) counted to at least 99 without assistance and without making more than three errors. Productive Counters had a median Final Highest Count of 99 (*M* = 96.4, *SD* = 9.5, *Range* = 49 to 99), which was 34 numbers higher than their median Initial Highest Count of 65. About a third of Non-Productive counters (*n*=15) received decade support (*Mean number of prompts* = 1.2, *SD* = 0.4, *Range* = 1 to 2), and the remaining children (*n*=34) initially made mid-decade errors and therefore did not receive decade prompts. Non-Productive Counters had a median Final Highest Count of 29 (*M* = 32.0, *SD* = 17.6, *Range* = 5 to 99), which was only 14 numbers past their median Initial Highest Count of 15. This difference is of course not surprising, since the ability to count at least 2 decades past the experimenter's first prompt was what defined the difference between Productive and Non-Productive children.

In summary, we found that between the ages of 4-and-a-half and 5-and-a-half, many children in our study exhibited evidence of having learned a productive decade+unit rule. Overall, the average age of Productive Counters was 5 years, 3 months. Though some of these

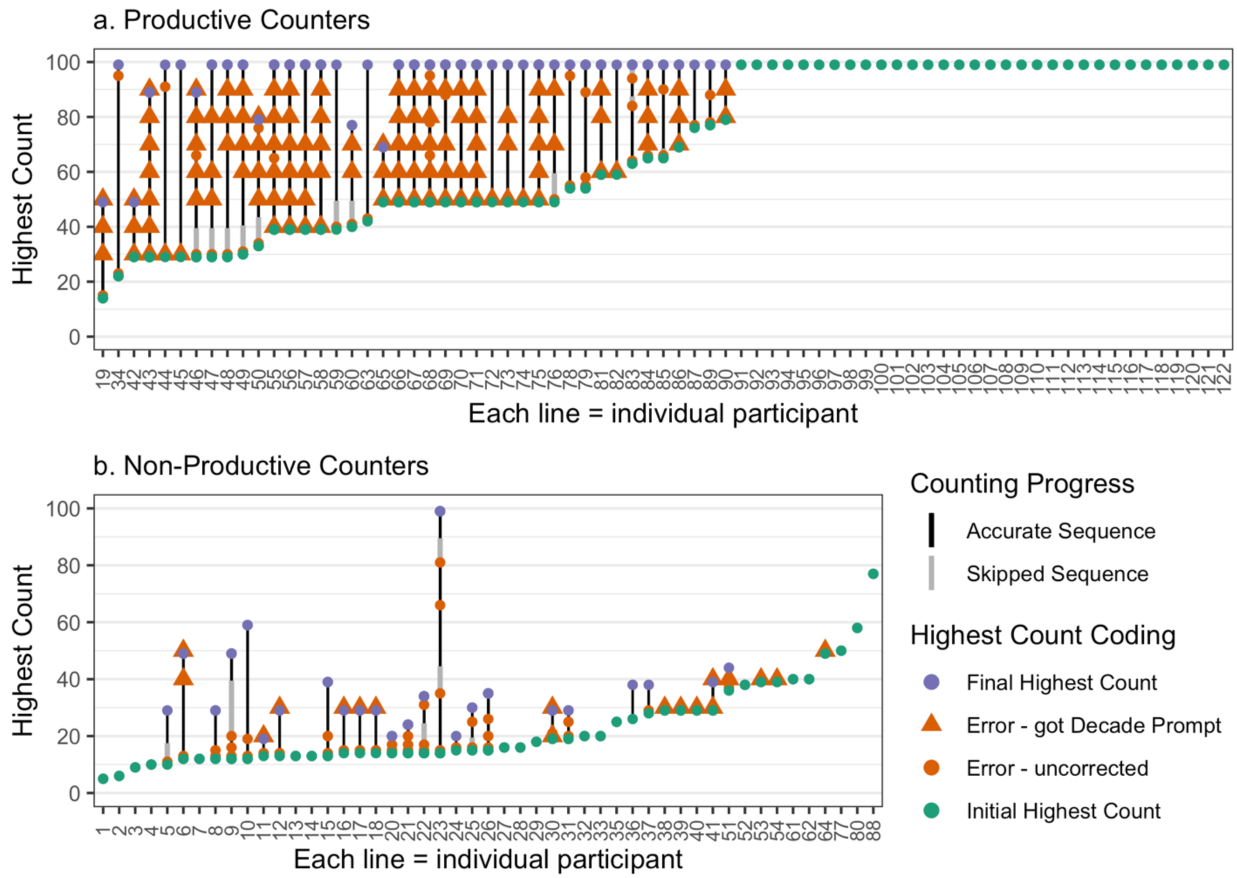

Figure 3. Highest count profiles for (a) Productive Counters and (b) Non-Productive Counters. Each line represents highest counts achieved by an individual child. Initial Highest Count (green dots) show highest number counted to before any error. Final Highest Count (purple dots) show highest number counted to as part of a consecutive 3-number sequence, allowing for up to 10 total errors and not more than three errors in a single decade. Errors include counting mistakes made on single numbers (orange dots and triangles) or over continuous intervals (orange lines). Decade Prompts (orange triangle) show decade terms provided by the experimenter after an error at a decade transition (e.g. stopping at 39). Black lines represent absolute gain from Initial to Final Highest Count.

children (*n*=41) were classified as Productive on the basis of having counted up to 99 with minimal errors, many received this classification because, upon stalling on a decade transition, they were able to recover once provided with a decade label, compatible with the use of a rule. These data not only provide evidence that children who stop on decade transitions do likely count using productive rules, but also suggests that a child's initial highest count may not provide the best measure of their mastery of counting. Even many children who had quite low initial counts (e.g., below 30) were able to recover when provided a decade prompt, suggesting that some children may memorize only a small subset of the count list (e.g., less than 30) before extracting a rule.

3.1.2 Productivity and Next Number performance. The Productivity classification, above, was one of two measures of decade+unit rule knowledge that we explored in this study. We also tested this using the Next Number task. Here, we asked how our two candidate measures of decade+unit rule knowledge were related to one another.

First, we found that Productive Counters (71% correct; *SD* = 27%) significantly outperformed Non-Productive Counters (28%; *SD* = 26%) on the Next Number task ($t(120) = -8.76$, $p < .001$). However, recall that Productive Counters included children who could count to 99 on their own (IHC ≥ 99) and those who could not (IHC < 99). As shown in Fig. 4, Productive Counters with IHC ≥ 99 performed close to ceiling on the Next Number task (*M* = 91%, *SD* = 15%), likely because all the items tested were within their familiar count sequence. To obtain a stronger, more conservative test of our hypothesis that Productive Counters could utilize a decade rule to succeed on this task, we excluded Productive Counters with IHC ≥ 99 from the following analyses. Additionally, because Productive Counters have higher Initial Highest Count and age than Non-Productive Counters, we deployed a regression strategy to include those two covariates.

Our second analysis thus used a mixed effects logistic regression predicting trial-level accuracy from Productivity, Initial Highest Count, age, and a Productivity by Initial Highest Count interaction (with random intercepts for subject and item magnitude). Although accuracy was twice as high among Productive Counters (IHC<99) (M=56%, SD=26%) compared to Non-Productive Counters (*M* = 28%, *SD* = 26%), this difference did not meet the threshold of

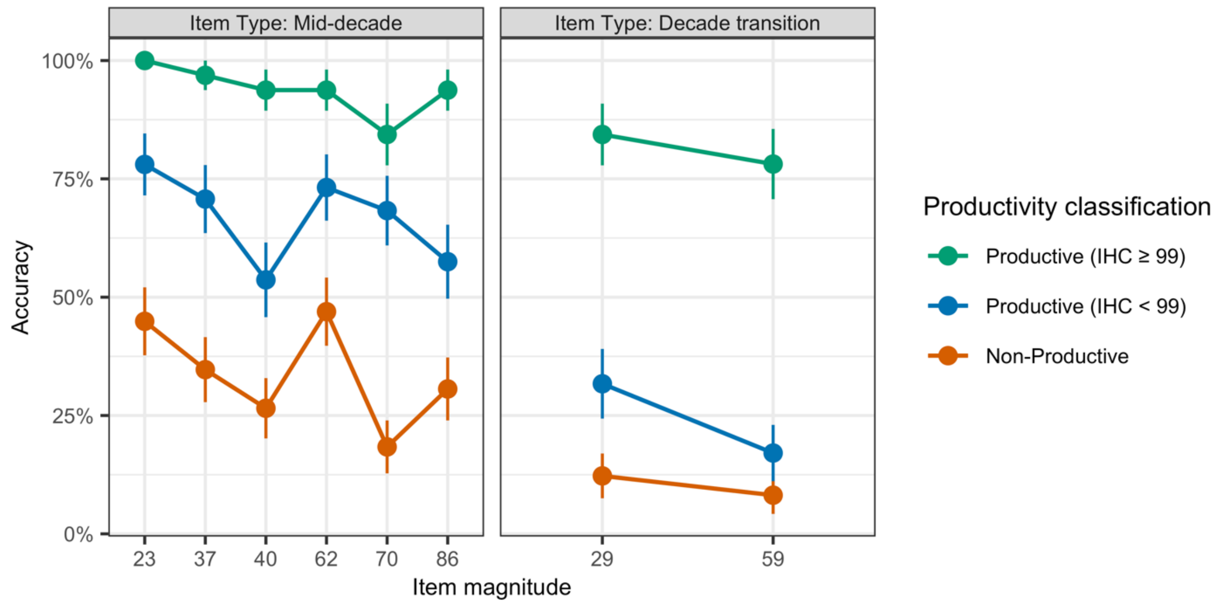

Figure 4. Average proportion of correct response on Next Number task by item and Decade Productivity status. Error bars indicate standard error about the mean.

statistical significance (β = 0.84, OR=2.31, 95% CI= [1.01, 5.30], $\chi^2(1) = 3.00$, $p = .08$), though there was a significant effect of Initial Highest Count (β = 0.90, OR=2.47, 95% CI= [1.61, 3.78], $\chi^2(1) = 14.9$, $p < .001$). Age was also not a significant predictor ($\chi^2(1) = .0001$, $p = .99$). However, these effects are qualified by a significant Productivity by Initial Highest Count interaction (β = -1.07, OR=.34, 95% CI= [.16, .76], $\chi^2(1) = 7.01$, $p = .008$). Inspection of simple slopes within Productivity group found that while Initial Highest Count positively predicted accuracy among Non-Productive Counters (β = 1.39, OR=4.01, 95% CI= [2.22, 7.24]), there was no effect for Productive Counters (β = 0.32, OR=1.38, 95% CI= [0.78, 2.43]).[6]

This second analysis suggests the Next Number task and children's ability to count-up from counting errors (i.e., our Productivity classification), capture different aspects of counting

[6] We obtain similar results when all Productive counters are included: there was a significant effect of Initial Highest Count, but not Productivity or Age. However, the interaction of Productivity and Initial Highest Count was no longer significant, likely because Productive Counters (IHC≥99) performed at ceiling. Full analysis results are reported in Supplemental Materials (B).

knowledge. Whereas Initial Highest Count was a strong predictor of Next Number performance for Non-Productive Counters, it was not related to accuracy for Productive Counters. This was not because Productive Counters had a limited range of Initial Highest Counts (indeed, some were as low as 29). Instead, one likely reason is that the Next Number and Productivity measures reflect performance on different ranges of the count sequence. Specifically, the Next Number task tests knowledge on both small and large numbers (e.g., from 23 to 86), whereas to be classified as Productive, a child might be tested with only relatively small numbers (e.g., counting up to 49 from an initial error on 29, see Fig. 3). Also, the Next Number task may place greater working memory demands on children – since it requires counting up from arbitrary points in the count list – potentially making a more strongly routinized count list more valuable, requiring less working memory resources for the retrieval of numbers.

We further explored how the two tasks are related by conducting a *post-hoc* analysis testing whether Productive Counters (IHC<99) show a selective advantage on the Next Number task for Mid-Decade items, where a productive decade rule ("*N*-ty-one, *N*-ty-two, …") would be most beneficial, compared to decade transitions, where this rule would be less beneficial. Including this Item Type variable (Mid-Decade or Decade Transition) significantly improved model fit ($\chi^2(1) = 12.4$, $p < .001$). This model found a significant main effect of Item Type ($\beta = 2.27$, OR=9.67, 95% CI = [4.03, 23.20], $\chi^2(1) = 25.79$, $p < .001$), with greater accuracy on Mid-Decade items ($M = 49\%$, $SD = 50\%$) than on Decade Transition items ($M = 17\%$, $SD = 37\%$). Adding a Productivity by Item type (Decade transition/Mid-decade) interaction did not significantly improve model fit ($\chi^2(1) = 0.29$, $p = .59$), indicating that both Productive Counters and Non-Productive Counters found mid-decade items easier than decade transition items.[7]

[7] Again, we obtain similar results when all participants are included in the analysis.

Next, we reasoned that, if a memorized count list is what allows children to generate successors on the Highest Count and Next Number tasks, then children should perform better on items within their Initial Highest Count than beyond it. If, instead, children have acquired a rule that generates the decade structure, then we might find an interaction between Productivity Group and Item Range, such that Productive Counters perform well both within and beyond their Initial Highest Count, while Non-Productive Counters are only able to generate successors within their Initial Highest Count. To test this prediction, we conducted a logistic mixed effects regression predicting trial-level accuracy from Productivity, Item Range (Within/Beyond IHC), Initial Highest Count, age, and the interaction of Productivity and Item Range, with random intercepts for subject and item magnitude. For this analysis, we excluded children with Initial Highest Count ≥ 99 because all numbers tested on this task would be within their counting range[8]. Model comparison by Likelihood Ratio Test found no significant main effect of either Productivity (p = .07) or Item Range (p = .89). Critically, there was a significant interaction effect of Productivity and Item Range (β = 1.16, OR= 3.20, 95% CI = [1.11, 9.21], $\chi^2(1) = 4.67$, p = .031). Planned contrasts indicate that performance on numbers within children's Initial Highest Count was similar for Productive Counters (IHC<99) (M = 61%) and Non-Productive Counters (M = 61%; p = .99 by t-test) whereas accuracy for numbers beyond their Initial Highest count was significantly greater among Productive Counters (M = 56%) than among Non-Productive Counters (M = 27%; $t(88) = 4.98$, $p < .001$; see Fig. 5).

Collectively, analyses contrasting children's highest count behaviors and their performance on the Next Number task suggest that these measures capture slightly different

[8] We do not perform this analysis including all participants, because Productive counters with Initial Highest Count ≥ 99 did not receive any trials outside their Initial Highest Count

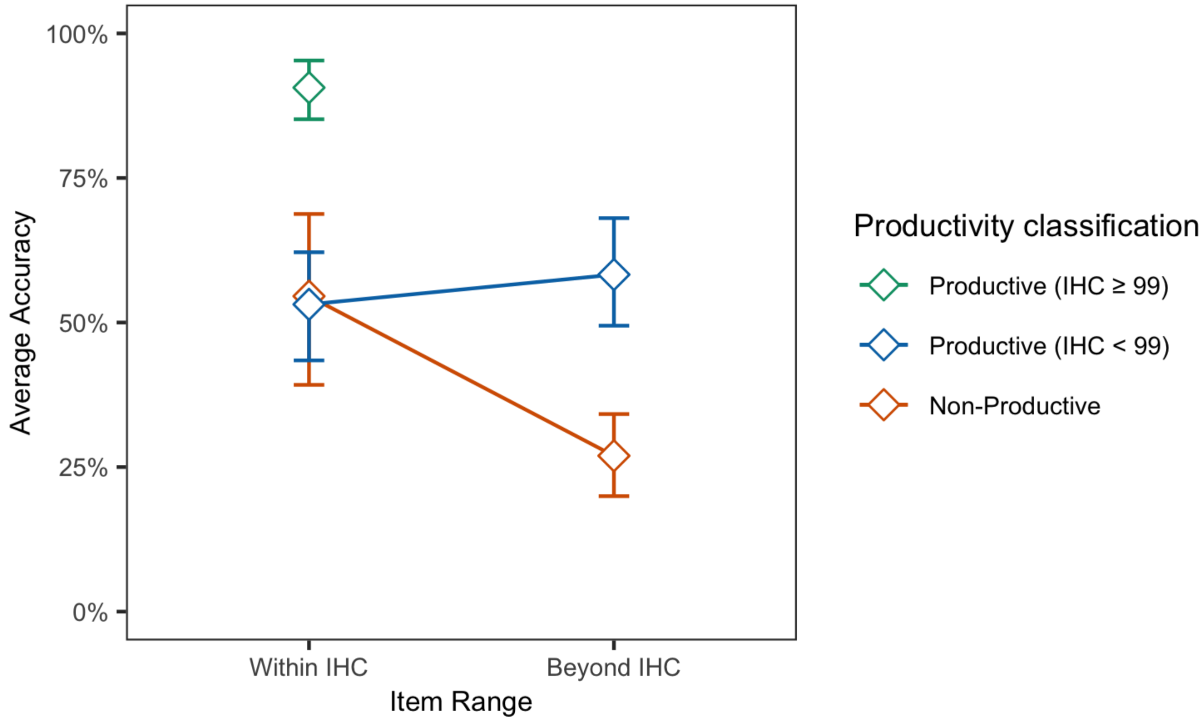

Figure 5. Average proportion correct for each participant on the Next Number task by Decade Productivity Status (between subjects) and Item Range (within subjects). Within IHC: correct answer is within a particular participant's Initial Highest Count; Beyond IHC: correct answer is greater than that participant's Initial Highest Count. Error bars show bootstrapped 95% confidence intervals about the mean. The largest number tested on this task was 86, thus, all trials were within IHC for productive counters with IHC ≥ 99.

phenomena. Both require children to draw on knowledge of the numeral sequence, but only the Next Number task requires children to generate number sequences from arbitrary points in the count list, without the benefit of the "momentum" afforded by the count routine for small and large numbers. Previous work finds that children's ability to count-up from an arbitrary position in the count list is significantly affected by an experimenter verbally rehearsing preceding numbers, and thereby providing momentum (e.g. Fuson, Richards, & Briars, 1982; Siegler & Robinson, 1982). In this sense, the Next Number task is a more difficult task, and perhaps a more stringent test of children's knowledge of the decade+unit rule.

3.2 Characterizing Infinity Knowledge

In the next set of analyses, we report children's successor function knowledge and beliefs about infinity. In the following section, we then probe how these are related to children's performance on the Highest Count and Next Number tasks.

We coded whether children thought every number has a successor ("Successor" knowledge) and separately, whether they thought numbers never end ("Endless" knowledge). Following previous work (Cheung et al., 2017), we also identified children who endorsed both beliefs as having "Full Infinity" knowledge (though we should note that, naturally, there is much more to fully grasping the notion of infinity than these two pieces of knowledge).

Table 2
Frequency of infinity knowledge in productive and non-productive counters

Classification	Non-Productive Counters (N = 49)	Productive Counters, IHC < 99 (N = 41)	Productive Counters, IHC ≥ 99 (N = 32)	Total (N = 122)
No Infinity Knowledge	33 (67%)	18 (44%)	8 (25%)	59 (48%)
Only Successor Knowledge	12 (24%)	10 (24%)	7 (22%)	29 (24%)
Only Endless Knowledge	1 (2%)	4 (10%)	6 (19%)	11 (9%)
Full Infinity Knowledge	3 (6%)	9 (22%)	11 (34%)	23 (19%)

Table 2 reports the frequency of these outcome classifications. About half of our sample (48%, n = 59) demonstrated no knowledge of infinity, claiming that there was a biggest number and that it was not possible to keep adding one. Non-Productive counters were more likely to fall in this category than Productive counters (67% of Non-Productive vs. 36% of Productive counters, $\chi^2(1) = 10.58$, $p = .001$). A smaller fraction of our sample (19%, n = 23) exhibited full knowledge of infinity, claiming that there was no biggest number and that we could always keep adding one. Productive counters were more likely to exhibit Full Infinity knowledge than Non-

Productive counters (27% vs. 6%, $\chi^2(1) = 7.34$, $p = .007$). The remaining children (33%, $n = 40$) had partial knowledge of infinity – claiming either that you could always add one ($n = 29$) or that there was no biggest number ($n = 11$), but not both. The distribution of infinity knowledge categories is comparable to previously reported results (e.g., Cheung et al., 2017; Evans, 1983; Hartnett & Gelman, 1998), although our sample consists of slightly younger participants.

3.3 Predictors of Infinity Knowledge

In this section we address how our measures of counting and productivity knowledge (Initial Highest Count, Productivity Group, Next Number task) were related to beliefs about infinity. Because the two components of infinity knowledge might develop independently, as indicated by the partial knowers who have one component knowledge but not the other, we conducted separate analyses to predict children's belief that every number has a successor (Successor Knowledge) and belief that numbers never end (Endless Knowledge). This allowed us to identify possible similarities or differences in relevant factors for acquiring each component of infinity knowledge. We also predicted children's status as Full Infinity Knowers, following Cheung et al. (2017).

Once again, our analyses excluded Productive Counters with IHC ≥ 99, so as to obtain a stronger, more conservative test of our hypothesis that Productive knowledge of counting might relate to beliefs about infinity.[9] Thus, we compared Productive Counters with IHC < 99 against Non-Productive Counters (total $N = 90$).

[9] Analyses incorporating the full sample are included in the Supplemental Materials (C). Those analyses qualitatively mirror the results presented here, though with some differences: None of the counting productivity measures were predictive of Successor knowledge. All three measures, however, were predictive of children's Endless knowledge, with Productivity group showing the largest effect. No single counting measure remained significant when entered into a full model predicting Endless knowledge, suggesting that the three measures explain overlapping variance. Finally, Next Number performance

For our initial models, predictors included Initial Highest Count, Productivity Group (Productive vs. Non-Productive Counters), and Next Number accuracy (proportion correct). Any predictor in these initial models that significantly predicted the outcome variable relative to a base model (with age as the only predictor) was added to a final model, to allow comparison among the predictors. Additionally, because Initial Highest Count was significantly correlated with both Productivity Group ($\chi 2(1) = 40.27$, $p < .001$) and Next Number accuracy ($r(88) = .59$, $p < .001$), whenever these latter two variables were entered in the final model we also included Initial Highest Count and its interaction in order to test for the role of decade+unit rule knowledge above and beyond rote counting ability. Models were constructed hierarchically, with model comparisons performed at every step using a Likelihood Ratio Test, and with models selected on the basis of a significant chi-squared statistic and reduced AIC value. For details about model fits and model comparison results, see Tables 3-5.

Interestingly, in predicting children's Successor Knowledge (Table 3), none of the three predictors explained a significant proportion of additional variance compared to the base model. In contrast, for models predicting children's possession of Endless Knowledge (Table 4), Productivity Group explained significant additional variance relative to the base model ($\chi^2(1) = 5.52$, $p = .019$; AIC = 84.9), though other measures of counting ability (Initial Highest Count, Next Number accuracy) did not explain additional variance when controlling for age. This final model thus included only Productivity Group and age as predictors, and estimated that Productive counters were more likely than Non-Productive Counters to have Endless Knowledge ($\beta = 1.63$, $p = .03$, OR = 5.08, 95% CI=[1.30,23.50]), while Age was not a significant predictor ($\beta = .02$, $p = 0.94$). The effect of Productivity Group remained significant when controlling for

explained significant additional variance in predicting Full Infinity Knowledge, but the coefficient was not significantly different from 0 ($p = .059$), suggesting overlapping variance with age.

Initial Highest Count, although these more complex models did not explain any additional variance. Finally, we constructed models predicting children's status as Full Infinity Knowers (Table 5). None of our counting measures (Initial Highest Count, Productivity group, or Next Number accuracy) improved model fit compared to the base model with only age as a predictor.

In summary, we found that children who were classified as Productive, based on the Highest Count task, were significantly more likely than Non-Productive children to believe that numbers never end. Other measures, such as Initial Highest Count and Next Number performance, were not as strongly related to infinity knowledge. In addition, none of our measures of counting knowledge were related to children's Successor Knowledge as measured by the infinity interview. Thus, in this study we find that one measure of children's productive counting rules predicts a belief that numbers are endless, but none of these measures predict the belief that it's always possible to add +1 to a number. This suggests that children may learn successor relations between numbers independent of learning productive counting rules – e.g., that this +1 rule describes the finite set of that they know, but that a morphological decade rule may explain why children believe that numbers never end – i.e., because number *words* can be productively generated.

Table 3

Regression models for predicting Successor Knowledge on the Infinity Interview (Participants IHC <99, N=90)

Models	Coefficient Estimates (ß)				Summary statistics		
	Age	IHC	Next Number accuracy	Productivity Group	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model							
Age	0.296				-58.75	121.51	0.027
Initial Models							
Age + IHC	0.598*	-0.520			-57.05	120.10	0.077
Age + Next Number accuracy	0.358		-0.155		-58.55	123.09	0.034
Age + Productivity Group	0.157			0.496	-58.32	122.64	0.040

Notes. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.
[a] Each initial model was compared against the base model.
* $p < 0.05$

Table 4

Regression models for predicting Endless Knowledge on the Infinity Interview (Participants IHC <99, N=90)

Models	Coefficient Estimates (ß)				Summary statistics		
	Age	IHC	Next Number accuracy	Productivity Group	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model							
Age	0.451				-42.22	88.44	0.049
Initial Models							
Age + IHC	0.261	0.349			-41.63	89.26	0.070
Age + Next Number accuracy	0.287		0.466		-41.04	88.07	0.090
Age + Productivity Group	0.025			1.625**	-39.46*	84.92	0.142
Final Models							
Age + Productivity Group + IHC	0.012	0.044		1.586**	-39.45	86.91	0.142
Age + Productivity Group * IHC	0.009	0.105		1.555**	-39.40	88.79	0.144

Notes. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.
[a] Each initial model was compared against the base model.
* $p < 0.05$, ** $p < 0.01$

Table 5

Regression models for predicting Full Infinity Knowledge on the Infinity Interview (Participants IHC <99, N=90)

Models	Coefficient Estimates (ß)				Summary statistics		
	Age	IHC	Next Number accuracy	Productivity Group	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model							
Age	0.527				-33.91	71.81	0.058
Initial Models							
Age + IHC	0.481	0.084			-33.88	73.76	0.059
Age + Next Number accuracy	0.417		0.307		-33.51	73.03	0.073
Age + Productivity Group	0.207			1.229	-32.71	71.43	0.104

Notes. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.

[a] Each initial model was compared against the base model.

4. Discussion

Given only finite experience with discrete quantities, number words, and counting, how do children learn that *every* natural number has a successor, and that numbers are endless? In this paper, we had two goals. First, we sought to characterize children's acquisition of productive morphological rules, and when this knowledge emerges in development. Second, we asked how such knowledge might be related to (a) their knowledge of the successor function (i.e., that it's possible to add +1 to any number), and (b) their beliefs regarding infinity (i.e., that numbers never end). Prior research suggests that how high children can count is related to their ability to identify successor relations for known numbers (e.g., Cheung et al., 2017), leading to the suggestion that counting experience causes children to notice the recursive base-10 structure of the count list, which in turn provides a basis for learning about successor relations and for generating unbounded number words (Barner, 2017; Cheung et al., 2017; Rule, Dechter, & Tenenbaum, 2015; Yang, 2016). Learning a rule that generates successive number words might lead children to the belief that all numbers have successors, and that numbers never end. Our study found multiple pieces of evidence that some 4- and 5-year-old children, but not others, use a productive rule when counting. Also, we found that Productive counters differed from Non-Productive counters with respect to their understanding of numerical infinity, though, interestingly, not their successor function knowledge, *per se*.

Several results provide evidence that some, but not all, 4- and 5-year-old children use a productive decade rule when counting (such that Productive Counters had a mean age of 5;3). First, when asked to count as high as they can, many children stopped at decade transitions, with about half making a decade-transition error as their first or second error. Whereas a memorized sequence predicts that errors should be randomly distributed over the count list, a decade rule predicts that they should occur disproportionately for irregular words that are not generated by a rule, such as decade labels (e.g., *twenty, thirty, fifty*). Furthermore, we found that when provided with decade prompts, many (65%) of the children who made decade transition errors could continue counting, often counting two or more decades further. Second, data from the Next Number task found that some children correctly generated successors for numbers beyond their Initial Highest Count, whereas some children did not, and could only name next numbers for numbers within their initial count. This, too, suggests that while some children's count list was purely memorized, unabetted by a decade rule, other children benefited from a rule that allowed them to identify next numbers on trials outside their familiar count routine. Third, we found some evidence that these two measures of productivity were related to one another, though imperfectly so. First, although Productive Counters had higher average accuracy on the Next Number task than Non-Productive Counters, the strongest predictor of Next Number performance was children's Initial Highest Count, not their Productivity classification. Second, whereas Productive and Non-Productive children performed similarly on the Next Number task for numbers within their Initial Highest Count, Productive children performed significantly better than Non-Productive children for numbers outside their Initial Highest Count, resulting in a significant interaction (and compatible with the use of a productive rule). The fact that Initial Highest Count was a stronger predictor of Next Number performance than Productivity suggests that, though in their own ways compelling measures of children's decade rule knowledge, these tasks draw on different constructs, perhaps because only the Next Number task requires children to count-up from arbitrary numbers without the benefit of momentum afforded by the count routine, and requires knowledge of both small and large numbers (Fuson et al., 1982; Siegler & Robinson, 1982).

In addition to characterizing several measures of counting productivity, our second goal was to explore how such measures might be related to children's intuitions about infinity. Overall, we found that children's belief that numbers never end was predicted by Productivity classification, but not other measures of counting proficiency, suggesting that children's ability to count-up from a decade label provided by the experimenter is the best predictor of whether they think numbers never end. Interestingly, we also found that no measures of counting proficiency were related to children's successor function knowledge, as measured by the infinity interview. Specifically, children's judgment that it's always possible to add 1 to a number didn't appear to be related to how high they could count, whether they could readily count-up from a decade label, or whether they could identify the next number in an arbitrary position in the count list. While previous studies have found that children with full infinity knowledge often have a highest count around 100 (e.g., Cheung et al., 2017; Hartnett & Gelman, 1998), the present study is the first to investigate how counting proficiency relates to successor and endless knowledge of infinity separately. This combination of results suggests that children's knowledge of successor relations may be acquired separately from their intuition that numbers never end, either because the successor function is not necessarily the basis by which children learn about infinity, or because successor function knowledge is initially defined over a finite set of numbers, and only later rendered fully recursive – e.g., by learning that recursive morphological rules of generate an infinite set of number words, and thus that the successor function can also be infinitely applied.

A starting point for this work was the observation in previous studies (Cheung et al., 2017) that children's acquisition of generalized successor function knowledge appears to be related to how high children can count. Critically, however, Cheung et al. only tested how counting abilities are related to children's reasoning about specific numbers, as measured by the

Unit task (Sarnecka & Carey, 2008), but not whether counting was related to the two beliefs about how numbers behave in general. Our work tested whether counting – specifically learning a productive decade rule – might explain more general intuitions regarding the successor function and infinity. We reasoned that counting might be related to such intuitions in two broad ways. First, it might be the case that, as children are increasingly exposed to numbers, they acquire more knowledge about how those numbers operate (including but not restricted to successor relations), which they may generalize to all numbers, without making a specific connection between learning morphological rules of counting and discovering that numbers are infinite. An alternative, however, is that counting abilities might relate to knowledge of infinity specifically because the morphological rules that govern counting provide rules for generating ever larger numbers. Such rules might provide the basis for the belief that numbers never end. That is, learning the morphological rules may allow children to reason that number words can be productively generated, and thus conclude that numbers are endless. On this view, counting may be separately related to the belief that every number has a successor and to the belief that numbers are endless.

Our data are compatible with this distinction between successor function and infinity knowledge. Children may learn item-based successor relations early on, and may even believe that all numbers have a successor, despite believing that only a finite number of numbers exist – i.e., they think the +1 function is bounded to a finite count list. About a quarter of our sample held such beliefs. Alternatively, children may have beliefs about numbers that support certain logical inferences, but not be aware of the entailments of these beliefs – e.g., that a fully recursive function necessitates that numbers never end. As evidence for this, Harnett and Gelman (1998; see also Hartnett, 1991) note that children's beliefs about whether or not numbers are infinite can change over the course of a single testing session, simply by asking them what the highest number is, and whether it's possible to add 1 to it repeatedly. For instance, some children in their study were able to recognize the inconsistency between their belief that there was a largest number and their belief that it is always possible to add one, and by the end of the testing session, had given up on either of these beliefs.

Our finding that children may acquire a productive counting rule by the time they have memorized a count list as short as 29 (see Figs. 2 and 3) presents a challenge to current computational models of number word learning. For example, Yang (2016) developed a "Tolerance Principle" which states that children invoke rules for explaining regularities in linguistic input when the number of exceptions or irregularities are below some threshold. According to his model, if there are N different linguistic tokens in the input, then a regular rule will be preferred only if the number of exceptions is below $N/\ln(N)$. In English, the early number words from 1-20 are exceptions to the decade rule, so the Tolerance Principle predicts that children would have to acquire a count list of at least length 72 before inducing a regular rule. Similar estimates in the 60-70 range were obtained by Rule, Dechter and Tenenbaum (2015) using a Bayesian architecture for inferring word to quantity mappings. However, our data suggest that some children can acquire a productive counting rule with much less data: the median Initial Highest Count of Productive Counters was only 49 (ignoring those who reached 99 on their own). Future work should reconcile these empirical findings within computational models of number word learning.

Future work might also explore how intuitions regarding infinity may come from sources beyond the count list. One possible source, for example, is experience with formal mathematical operations like addition, or the recursive addition of zeros in Arabic notation, which could help children appreciate ways of generating endlessly more quantities. For example, Singer and Voica (2008) describe how a group of 5th and 6th-graders explained that rational numbers are infinite, because infinitely many digits could be added after the decimal point. Another possible source of intuitions is by analogy to quantities other than number, such as space, time, or geometry. One early study by Evans (1983) found that knowledge about infinity in number, time, and space developed similarly between Kindergarten and 3rd Grade (e.g., from bounded to unbounded), but this research did not test if such beliefs were correlated within individual children (see Hartnett & Gelman, 1998, for similar findings). Current studies in our labs are exploring this possibility. More recently, Smith, Solomon and Carey (2005) found that 1st and 2nd graders' intuitions matter being infinitely divisible preceded their intuitions about the infinite divisibility of rational numbers. Numbers can be infinite in many ways (Monaghan, 2001); correspondingly, there is much room for future research into children's intuitions about infinity in different conceptual domains.

Much remains to be discovered about how children acquire rule-like knowledge of the count sequence and such knowledge helps shape beliefs about the logical structure of natural numbers. One limitation of the present work is that all tasks used real, attested, numbers. Although our productivity analyses were meant to use errors as an indication of rule use, it remains possible that children's errors on decade terms were nevertheless affected by differences in rote memory for decade terms vs. other count words (e.g., that somehow children's memory trace of "forty" is weaker than their memory trace of "forty-one" or "forty-two"). Any study that tests children's knowledge of attested number words faces this type of challenge. A stricter - and possibly more onerous - test of understanding the decade+unit rule might involve unfamiliar or made-up numbers. For instance, a child with productive knowledge might be able to judge that the successor of “a billion one” is “a billion two” and that the successor of “daxy-five” is “daxy-six”. Recent work testing this prediction has found that children’s ability to count up from these less familiar or novel numbers is correlated with the measures of productivity used here, and is a strong predictor of their performance on a task assessing successor function knowledge (Schneider, Sullivan, Guo, & Barner, *in prep*).

In addition to the factors we tested, it is likely that other domain general capacities play a role in children's ability to succeed at our experimental tasks, including the ability to hold numbers in working memory, and the more general ability to extract patterns - or rules - from linguistic data. Such differences may help explain why some children with low Initial Highest Counts were able to count higher using provided decade labels, whereas other children with higher Initial Highest Counts were not. However, previous studies suggest that domain general limits are unlikely to fully explain individual differences in count list proficiency. In one cross-cultural study of successor function knowledge, English-speaking US children outperformed children in India learning morphologically complex count lists in Hindi and Gujarati, even when controlling for working memory, age, and Initial Highest Count (Schneider et al, 2020). Thus, while domain-general cognitive abilities are likely recruited in these tasks, they are importantly deployed over language-specific inputs that a child receives.

In conclusion, we suggest that sometime around 5 years of age, children learn to generate number words beyond their spontaneous counting range by implementing a recursive base-10 rule defined over their verbal count list. This insight may support an inductive inference over their acquired verbal count sequence, which facilitates conceptual insights into the infinite nature of the natural numbers. Consistent with recommendations from early childhood educators (Frye et al., 2013), the present results demonstrate another way that forward counting ability (i.e., counting up from one) may provide children with a useful foundation for learning more abstract natural number concepts such as cardinality, the successor function, and arithmetic. Our results suggest that understanding the logic of natural numbers is closely related to understanding the syntactic logic of the verbal numerals, suggesting one route by which learning language can impact learning about number concepts.

References

Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(46), 18448–18453.

Barner, D. (2017). Language, procedures, and the non-perceptual origin of number word meanings. *Journal of Child Language*, *44*.

Barner, D., Libenson, A., Cheung, P., & Takasaki, M. (2009). Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *Journal of Experimental Child Psychology*, *103*(4), 421–440.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4, *67*(1). doi:10.18637/jss.v067.i01

Carey, S. (2004). Bootstrapping & the origin of concepts. *Daedalus*, *133*(1), 59–68.

Cheung, P., Dale, M., & Le Corre, M. (2016). A cross-linguistic investigation on the acquisition of complex numerals. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2711–2716). Austin, TX: Cognitive Science Society.

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.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction? *Cognition*, *123*(1), 162–173.

Decock, L. (2008) The conceptual basis of numerical abilities: one-to-one correspondence versus the successor relation. *Philos. Psychol.* 21, 459–473

Dedekind, R. (1888/1963). *Essays on the Theory of Numbers*. (W.W. Beman, Trans.). Dover.

Dowker, A., Bala, S., & Lloyd, D. (2008). Linguistic Influences on Mathematical Development: How

Important Is the Transparency of the Counting System? *Philosophical Psychology*, *21*(4), 523–538.

Evans, D. W. (1983). *Understanding zero and infinity in the early school years*. (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, PA.

Evans, D. W., & Gelman, R. (1982). Understanding infinity: A beginning inquiry. Unpublished manuscript.

Frye, D., Baroody, A. J., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). *Teaching math to young children: A practice guide (NCEE 2014-4005)*. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://whatworks.ed.gov

Fuson, K. C., Richards, J., & Briars, D. J. (1982). The acquisition and elaboration of the number word sequence. In *Progress in cognitive development research: Vol. 1. Children's logical and mathematical cognition,* pp.33–92.

Gelman, R., & Gallistel, C. R. (1978). *The child's concept of number*. Cambridge: Harvard University Press.

Gentner, D. (2010). Bootstrapping the Mind: Analogical Processes and Symbol Systems. *Cognitive Science*, *34*(5), 752–775.

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.

Grotenhuis, M., Pelzer, B., Eisinga, R., Nieuwenhuis, R., Schmidt-Catran, A., & Konig, R. (2017). When size matters: advantages of weighted effect coding in observational studies. *International Journal of Public Health*, *62*(1), 163–167.

Hartnett, P. M. (1991). *The development of mathematical insight: From one, two, three to infinity*. (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, PA

Hartnett, P. M., & Gelman, R. (1998). Early understandings of numbers: paths or barriers to the construction of new understandings? *Learning and Instruction*, *8*(4), 341–374.

Humboldt, W. von. (1836/1999). *On Language*. (P. Heath, Trans.). Cambridge: Cambridge University

Press.

Hurford, J. R. (1975). *The Linguistic Theory of Numerals*. Cambridge Studies in Linguistics (Vol. 16). Cambridge University Press.

Kaminski, J. (2015). Young children's understanding of the successor function. *Proceedings of the 37th Annual Conference of the Cognitive Science Society Meeting* (pp. 1045-1050). Pasadena, CA: Cognitive Science Society.

Le Corre, M., Li, P., Huang, B. H., Jia, G., & Carey, S. (2016). Numerical morphology supports early number word learning: Evidence from a comparison of young Mandarin and English learners. *Cognitive Psychology*, *88*, 162–186

Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence. *Psychological Science*, *6*(1), 56–60.

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

Monaghan, J. (2001). Young peoples' ideas of infinity. *Educational Studies in Mathematics*, *48*, 239–257.

Peirce, C. S. (1881). On the Logic of Number. *American Journal of Mathematics*, 4.

Piantadosi, S. T., Jara-Ettinger, J., & Gibson, E. (2014). Children's learning of number words in an indigenous farming-foraging group. *Developmental Science*, *17*(4), 553–563.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. Trends in Cognitive Sciences, 6(11), 456-463.

R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org

Rule, J., Dechter, E., & Tenenbaum, J. B. (2015). Representing and learning a large system of number concepts with latent predicate networks. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, 2051–2056.

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: what children must learn and when they learn it. *Cognition*, *108*(3), 662–674.

Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., & Yudovina, Y. B. (2007). From grammatical number to exact numbers: Early meanings of 'one', 'two', and 'three' in English, Russian, and Japanese. *Cognitive Psychology*, *55*(2), 136–168.

Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology*, *103*(3), 325–337.

Schaeffer, B., Eggleston, V. H., & Scott, J. L. (1974). Number development in young children. *Cognitive Psychology*, *6*(3), 357–379.

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, 101263. https://doi.org/10.1016/j.cogpsych.2019.101263

Schneider, R.M., Sullivan, J., Guo, K., and Barner, D. (in prep) What counts? Sources of knowledge in children's acquisition of the successor function

Sella, F. & Lucangeli, D. (2020). The knowledge of the preceding number reveals a mature understanding of the number sequence. *Cognition*, *194*.

Siegler, R. S., & Robinson, M. (1982). The development of numerical understandings. *Advances in Child Development and Behavior*, *16*, 279–312.

Singer, F. M., & Voica, C. (2008). Between perception and intuition: Learning about infinity. *Journal of Mathematical Behavior*, *27*(3), 188–205.

Smith, C. L., Solomon, G. E. A., & Carey, S. (2005). Never getting to zero: Elementary school students' understanding of the infinite divisibility of number and matter. *Cognitive Psychology*, *51*(2), 101–140.

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.

von Leibniz, G. W. (1996). *New essays on human understanding*. (P. Remnant & J. Bennett, Trans.). Cambridge: Cambridge University Press.

Wagner, K., Kimura, K., Cheung, P., & Barner, D. (2015). Why is number word learning hard? Evidence from bilingual learners. *Cognitive Psychology*, *83*, 1–21.

Wright, C. (1983) Frege's Conception of Numbers as Objects. Aberdeen University Press

Wright, R. J. (1994). A study of the numerical development of 5-year-olds and 6-year-olds. *Educational Studies in Mathematics*, *26*(1), 25–44.

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

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

Yang, C. (2016, April). The Linguistic Origin of the Next Number. Retrieved from https://ling.auf.net/lingbuzz/003824

Here we provide more detailed information about (A) the Infinity Interview and (B-D) report additional analyses exploring the robustness of our results to various cutoff criteria.

Our materials, data, and analysis code are available in a repository on the Open Science Framework: https://osf.io/z6ky3/

Table of Contents

A: Infinity Interview

We describe a detailed coding scheme for classifying responses to the Infinity Interview. We first provide coding rules (A1), followed by an example transcript for each of the four possible Infinity knowledge classifications (A2-A5). The Infinity Interview was described in the main text (section 2.2.3) with results presented in section 3.2.

A1: Detailed Coding scheme

The Infinity Interview consisted of 6 questions presented in the same order:

1. *"What is the biggest number you can think about?"* If the child did not answer, the experimenter probed them by asking how high they could count.
2. *"Is that the biggest number there could ever be?"*
    a. If yes, move on.
    b. If no, "*Can you think of a bigger number? Is that the biggest number there could ever be?"* The experimenter repeated this exchange up to 4 times or until the child affirmed that they had produced the biggest number.
3. "*If I keep counting, will I ever get to the end of numbers, or do numbers go on forever? Why?*"
4. "*If we thought of a really big number, could we always add to it and make it even bigger, or is there a number so big we couldn't add any more? Why / Why not?*"
5. "*You said the biggest number you know is X. Tell me, is it possible to add one to X, or is X the biggest number possible? Why?*" For this question, X was the largest number the child had stated in the entire testing session.
6. "*Could I keep adding one? Why / Why not?*"
    a. If yes, "*What would happen if I kept adding one?*"

We coded successor and endless knowledge of infinity separately, as two binary variables.

To have endless knowledge of infinity, participants had to respond that there was no biggest number (Q1, Q2, Q5) or have provided four responses to the request to provide a bigger number (Q2). Participants who claimed to have provided the biggest number at some point in the interview (Q1, Q2, Q5) or said that numbers end (Q3) were classified as not exhibiting Endless knowledge of infinity.

To have successor knowledge of infinity, participants had to respond that it was possible to add 1 to some big number (Q4 and Q5) or claim that it was possible to keep adding one to that big number (Q6).

A2: Infinity transcript for a child classified as having no knowledge of infinity

E = Experimenter
C = Child (Age: 4 years 1 month)

E: What is the biggest number you can think about?
C: One hundred.
E: Is that the biggest number there could ever be?
C: Yes.
E: If I keep counting, will I ever get to the end of numbers, or do numbers go on forever?
C: End.
E: Why?
C: Long and get to one hundred.
E: If we thought of a really big number, could we always add to it and make it even bigger, or is there a number so big we couldn't add any more?
C: No.
E: Why?
C: Already so big.
E: So you said that the biggest number you know is one hundred. Is it possible to add 1 to one hundred, or is one hundred the biggest number possible?
C: Yes.
E: Why?
C: There's also a hundred seventy.
E: Could I keep adding 1?
C: No.
E: Why?
C: It already has one.

A3: Infinity transcript for a child classified as having only Successor knowledge of infinity

E = Experimenter
C = Child (Age: 5 years 6 months)

E: What is the biggest number you can think about?
C: Sixty one.
E: Is that the biggest number there could ever be?
C: No, one hundred fifty nine sixty.
E: Is that the biggest number there could ever be?
C: I don't know.
E: Well, can you think of a bigger number?
C: Eighteen ninety sixty one.
E: Is that the biggest number there could ever be?
C: No.
E: Can you think of a bigger number?
C: No.
E: If I keep counting, will I ever get to the end of numbers, or do numbers go on forever?
C: Forever.
E: Why?
C: End.
E: Why?
C: actually, you won't (get to the end of numbers) my friend can count to 100 and after and I learnt to count to 100 but not really.
E: Do you think she could count forever?
C: I don't know.
E: If we thought of a really big number, could we always add to it and make it even bigger, or is there a number so big we couldn't add any more?
C: Yes.
E: Why?
C: It makes it bigger.
E: So you said that the biggest number you know is eighteen ninety sixty one. Is it possible to add 1 to eighteen ninety sixty one, or is eighteen ninety sixty one the biggest number possible?
C: Yes.
E: Why?
C: Eighty nine.
E: Could I keep adding 1?
C: Yes
E: Why?
C: It makes it bigger and bigger and bigger.
E: What would happen if I kept adding 1?
C: It will get too big.
E: Do we have to stop or could we keep adding 1?
C: Stop.

A4: Infinity transcript for a child classified as having only Endless knowledge of infinity

E = Experimenter
C = Child (Age: 4 years 9 months)

E: What is the biggest number you can think about?
C: One hundred.
E: Is that the biggest number there could ever be?
C: No, three hundred.
E: Is that the biggest number there could ever be?
C: No, four hundred.
E: Is that the biggest number there could ever be?
C: Yes.
E: If I keep counting, will I ever get to the end of numbers, or do numbers go on forever?
C: Forever.
E: Why?
C: There's a lot of numbers and they never end.
E: If we thought of a really big number, could we always add to it and make it even bigger, or is there a number so big we couldn't add any more?
C: Couldn't add more.
E: Why?
C: Because we couldn't add nine hundred to it.
E: So you said that the biggest number you know is four hundred. Is it possible to add 1 to four hundred, or is four hundred the biggest number possible?
C: No.
E: Why?
C: Take a long time.
E: Could I keep adding 1?
C: I don't know.

A5: Infinity transcript for a child classified as having Full knowledge of infinity

E = Experimenter
C = Child (Age: 5 years 10 months)

E: What is the biggest number you can think about?
C: A Jillion.
E: Is that the biggest number there could ever be?
C: No, infinity.
E: If I keep counting, will I ever get to the end of numbers, or do numbers go on forever?
C: Forever.
E: Why?
C: There's lots of numbers in the world.
E: If we thought of a really big number, could we always add to it and make it even bigger, or is there a number so big we couldn't add any more?
C: Yes.
E: Why?
C: Every number has a partner. Like 5 goes with 10s, twenties goes thirties, forties goes fifties, and sixties goes seventies, eighty goes ninety, one hundred goes one hundred and one.
E: So you said that the biggest number you know is infinity. Is it possible to add 1 to infinity, or is infinity the biggest number possible?
C: Yes.
E: Why?
C: I can if I want to.
E: Could I keep adding 1?
C: Yes.
E: What would happen if I kept adding 1?
C: I'd get to infinity, it'll be a million and one.

B: Analysis of Productivity and Next Number accuracy with all participants

This analysis is referenced in Section 3.1.2 of the main text (Footnote 6 and 7). In the main text, we reported analyses excluding participants with an Initial Highest Count of 99. Here, we compare Non-Productive Counters and Productive Counters in terms of their accuracy on the Next Number task when all participants are included.

A mixed effects logistic regression predicting item-level accuracy from age, Productivity, Initial Highest Count and Productivity by IHC interaction (with random intercepts for subject and item magnitude), found no significant effect of Productivity ($\chi^2(1) = 2.43$, $p = 0.12$) or Age ($\chi^2(1) < 1$, $p = 0.93$), or Productivity by IHC interaction ($\chi^2(1) = 2.45$, $p = .12$). However, there was a significant effect of Initial Highest Count ($\beta = 1.92$, OR=6.83, 95% CI=[3.99, 11.7], $\chi^2(1) = 53.79$, $p < .001$).

Next, we conducted a post-hoc analysis testing whether Productive counters may show a selective advantage on the Next Number task for mid-decade items but not decade transition items. This analysis found a significant main effect of Item Type ($\beta = 2.10$, OR=8.18, 95% CI=[3.46, 19.3], $\chi^2(1) = 22.96$, $p < .001$), with greater accuracy on Mid-Decade items ($M = 61\%$, $SD = 49\%$) than on Decade Transition items ($M = 34\%$, $SD = 47\%$). Adding a Productivity by Item type (Decade transition/Mid-decade) interaction did not significantly improve model fit ($\chi^2(1) < 1$, $p = .99$), indicating that both Productive Counters and Non-Productive Counters found mid-decade items easier than decade transition items. This result suggests that both Productive and Non-Productive Counters have more difficulty generating decade terms than generating successive numbers within the same decade.

C: Analysis of Infinity knowledge with all participants

These analyses are referenced in Section 3.3 of the main text (Footnote 11). In the main text, we exclude from analysis Productive Counters who had counted to 99 on their own without error. That allows for a more conservative test of the hypothesis that Productive knowledge of counting might relate to beliefs about the successor function and infinity. Here, we provide analysis results when all participants are included, and note any qualitative differences from the analysis in the main text.

We conduct separate analyses to predict children's belief that every number has a successor (C1: Successor Knowledge), their belief that numbers never end (C2: Endless Knowledge), and children's status as Full Infinity Knowers (C3: Full Knowledge). The analysis approach is reported in the main text; we constructed logistic regression models in a hierarchical fashion using Likelihood Ratio Tests to evaluate the contributions of additional variables, and selected final models based on a significant chi-squared statistic and reduced AIC value.

Note that, because Initial Highest Count was significantly correlated with both Productivity Group ($\chi 2(1) = 80.45$ $p < .001$) and Next Number accuracy ($r(88) = .77$, $p < .001$), whenever these latter two variables were entered in the final model we also included Initial Highest Count and its interaction in order to test for the role of decade+unit rule knowledge above and beyond rote counting ability.

C1: Regression analyses predicting Successor Knowledge of Infinity

(See Table S1). Similar to analyses with the partial sample, initial models predicting Successor Knowledge of Infinity found that none of the three predictors explained a significant proportion of additional variance compared to the base model.

C2: Regression analyses predicting Endless Knowledge of Infinity

(See Table S2). For models predicting children's possession of Endless knowledge, all three counting measures explained significant additional variance relative to the base model. Controlling for age, children were more likely to have Endless knowledge if they were Productive counters ($\beta = 1.70$, OR = 5.46, 95% CI = [1.69, 21.58], $\chi^2(1) = 8.38$, $p = .004$), had greater accuracy on the Next Number task ($\beta = 0.69$, OR = 2.00, 95% CI = [1.21, 3.46], $\chi^2(1) = 7.52$, $p = .006$), or had greater Initial Highest Counts ($\beta = 0.67$, OR = 1.96, 95% CI = [1.22, 3.20], $\chi^2(1) = 7.81$, $p = .005$).

To evaluate the relative contribution of each predictor, we constructed a full model which included all three predictors while controlling for age. This model explained significant additional variance compared to the base model with only age ($\chi^2(3) = 11.31$, $p = 0.01$), but did not improve model fit relative to any of the initial models. In addition, none of the predictor coefficients in this full model were significantly different from zero, suggesting that these measures of counting ability may explain overlapping variance in predicting Endless Knowledge among our sample.

C3: Regression analyses predicting Full Knowledge of Infinity

(See Table S3). For initial models predicting Full Infinity knowledge, accuracy on the Next Number task explained significant additional variance relative to the base model ($\chi^2(1) = 3.85$, $p = .0496$). However, the model estimated coefficient for Next Number accuracy did not meet the threshold for significance ($\beta = 0.57$, $p = 0.059$, OR = 1.77, 95% CI = 1.00, 3.34]).

Table S1

Regression models for predicting Successor knowledge on the Infinity Interview (All participants, N=122)

Models	Coefficient Estimates (ß)				Summary statistics		
	Age	IHC	NN	Productivity	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model							
Age	0.342				-81.54	167.08	0.037
Initial Models							
Age + IHC	0.304	0.074			-81.48	168.96	0.038
Age + NN	0.286		0.126		-81.36	168.72	0.041
Age + Productivity	0.163			0.661	-80.51	167.03	0.059

Notes. IHC: Initial Highest Count; NN: Next Number accuracy; Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.

[a] Each initial model was compared against the base model.

Table S2

Regression models for predicting Endless knowledge on the Infinity Interview (All participants, N=122)

Models	Coefficient Estimates (ß)						Summary statistics		
	Age	IHC	NN	Productivity	IHC * NN	IHC * Productivity	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model									
Age	0.707**						-66.66	137.31	0.125
Initial Models									
Age + IHC	0.403	0.675**					-62.75**	131.50	0.207
Age + NN	0.448		0.693**				-62.89**	131.79	0.204
Age + Productivity	0.329			1.698**			-62.46**	130.93	0.212
Additional Models									
Age + IHC * NN	0.392	0.364	0.402		.066		-62.14	134.29	0.219
Age + IHC * Productivity	0.268	0.353		1.320		0.206	-61.30	132.60	0.236
Full Model									
Age + IHC + NN + Productivity	0.251	0.251	0.291	1.092			-61.00*	131.99	0.242

Notes. IHC: Initial Highest Count; NN: Next Number accuracy. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.

[a] Using Likelihood Ratio Tests, each initial model and the full model was compared against the base model. Each additional model was compared to the corresponding initial model without IHC.

* p < 0.05, ** p < 0.01

Table S3

Regression models for predicting Full Infinity knowledge on the Infinity Interview (All participants, N = 122)

Models	Coefficient Estimates (ß)					Summary statistics		
	Age	IHC	NN	Productivity	IHC * NN	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model								
Age	0.752**					-54.36	112.71	0.120
Initial Models								
Age + IHC	0.560	0.400				-53.31	112.62	0.145
Age + NN	0.532		0.573			-52.43*	110.86	0.166
Age + Productivity	0.480			1.219		-52.78	111.57	0.158
Additional Models								
Age + IHC * NN	0.600*	-0.251	0.644		0.365	-51.93	113.87	0.178

Notes. IHC: Initial Highest Count; NN: Next Number accuracy. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.

[a] Each initial model was compared against the base model.

D: Reanalysis using a more conservative Productivity Classification

In this section we describe a stricter coding scheme for classifying participants' Productivity status based on the Highest Count Task (originally described in section 2.2.1 of manuscript).

These re-analyses led to the same conclusions as reported in the main text regarding how Productivity influences Highest Count performance (main text, section 3.1), Next Number performance (main text, section 3.2), and Successor Knowledge of Infinity (main text, section 3.2). However, there was a different result for regression analysis predicting Endless knowledge and Full knowledge of infinity. First, the original findings reported that Productivity group significantly predicted Endless knowledge; this effect did not replicate using the stricter definition. Second, the original findings did not yield any significant predictors of Full infinity knowledge when controlling for age; however, the stricter Productivity definition significantly predicts Full infinity knowledge, even when controlling for age or Initial Highest Count. Given that Full infinity knowledge requires Endless knowledge, these results generally converge.

In summary, these analyses indicate three robust findings: (1) there exists significant individual differences in knowledge of the decade+unit rule, which affects performance on counting and Next Number tasks; (2) knowledge of this productive rule does not predict successor function knowledge (i.e. the belief that we can always add one); (3) Productivity does predict infinity understanding (i.e. the belief that numbers never end), although particular regression results may depend on how strictly we define productivity and infinity knowledge (i.e. Endless knowledge vs. Endless+Successor knowledge).

D1: Relationship of Productivity to Highest Count and age

In this stricter definition, we allow for only one error when classifying participants' Productivity Status. Thus, to be Productive, participants have to count to 99 on their own with at most one error ("Productive Counters, IHC ≥ 99") or count past a Decade-Change Error by at least 2 decades, with at most one error in those 2 decades ("Productive Counters, IHC < 99"). This stricter definition reclassifies 6 participants as Non-Productive counters instead of Productive Counters. Their subject IDs (in main manuscript, Fig. 3) are: 34, 78, 79, 83, 85, 89. Below, we replicate analyses reported in the main manuscript (Section 3.1).

Productive Counters had, on average, a higher Initial Highest Count (*Mean* = 70.4, *SD* = 29.4, *Median* = 76) than Non-Productive Counters (*Mean* = 26.3, *SD* = 18.4, *Median* = 18). This difference remained if we considered only Productive Counters with Initial Highest Count below 99 (*Mean* = 44.3, *SD* = 14.3, *Median* = 49).

Productive Counters also made greater improvements past their Initial Highest Counts. Non-Productive Counters had a median Final Highest Count of 29 (*M* = 39.3, *SD* = 26.8, *Range* = 5 to 99), which was only 11 numbers past their median Initial Highest Count of 18. In contrast, Productive Counters had a median Final Highest Count of 99 (*M* = 96.1, *SD* = 9.9, *Range* = 49 to 99), which was 23 numbers higher than their median Initial Highest Count of 76. When considering only Productive Counters (IHC<99), the improvement between initial and final highest counts is even greater: their median Final Highest Count was 99, which was 50 numbers higher than their median Initial Highest Count of 49.

Table S4: Distribution of Age and Highest Count measures by Productivity

Classification	**Variable**	*n*	*M*	*SD*	*Median*	*Min*	*Max*
Original definition							
Nonproductive	Age (years)	49	4.60	0.42	4.49	4.00	5.61
	IHC	49	22.63	14.86	15	5	77
	FHC	49	32.00	17.62	29	5	99
	DCE	19	30.58	8.34	29	19	49
Productive	Age (years)	41	5.22	0.48	5.24	4.25	5.99
(IHC < 99)	IHC	41	46.02	15.34	49	14	79
	FHC	41	94.32	12.33	99	49	99
	DCE	34	47.82	15.33	49	29	89
Stricter definition							
Nonproductive	Age (years)	55	4.65	0.43	4.56	4.00	5.61
	IHC	55	26.25	18.37	18	5	77
	FHC	55	39.31	26.84	29	5	99
	DCE	20	33.50	15.38	29	19	89
Productive	Age (years)	35	5.26	0.49	5.35	4.25	5.99
(IHC < 99)	IHC	35	44.34	14.35	49	14	79
	FHC	35	93.51	13.21	99	49	99
	DCE	33	46.58	13.70	49	29	79

D2: Relationship of Productivity to Next Number performance

We replicate the effect of Productivity on Next Number performance: Productive Counters (72% correct; *SD* = 27%) significantly outperformed Non-Productive Counters (31%; *SD* = 27%) on the Next Number task ($t(120) = -8.13$, $p < .001$). Below, we conduct more conservative analyses considering just Productive Counters (IHC<99) and Non-Productive Counters. We use mixed effects logistic regression to predict trial-level accuracy from predictors of interest, with random effects for subject and item magnitude.

The finding that Initial Highest Count predicts Next Number accuracy for Non-Productive Counters, but not Productive Counters, is robust. We constructed a model predicting accuracy from Productivity, Initial Highest Count, age, and a Productivity by Initial Highest Count interaction and age. There was no significant main effect of Productivity ($\beta = 0.90$, $\chi^2(1) = 2.53$, $p = 0.11$) or Age ($\chi^2(1) = 0.026$, $p = .87$), but there was a significant main effect of Initial Highest Count (β = 0.85, OR=2.34, 95% CI=[1.56, 3.51]; $\chi^2(1) = 20.9$, $p < .001$). However, these effects are qualified by a significant Productivity by Initial Highest Count interaction ($\beta = -0.91$, OR=.40, 95% CI=[.19, .88], $\chi^2(1) = 5.20$, $p = 0.023$). Inspection of simple slopes within Productivity group found that while Initial Highest Count positively predicted accuracy among Non-Productive Counters ($\beta = 1.20$, OR=3.32, 95% CI=[2.09, 5.27]), there was no effect for Productive Counters ($\beta = 0.29$, OR=1.34, 95% CI=[0.68, 2.64]).

Next, we evaluated how productivity classification might interact with various item-level covariates. In the main text, a *post-hoc* analysis testing whether Productive Counters did not find a selective advantage on the Next Number task for Mid-Decade items compared to decade transitions. This model predicted accuracy from Productivity, Item Type, Productivity by Item Type interaction, and Initial Highest Count. Using a stricter definition of Productivity, we replicated this finding: there was no significant interaction of Productivity and Item Type ($\chi^2(1) = .67$, $p = .41$). Instead, there was a main effect of Item type ($\beta = 2.27$, OR=9.70, 95% CI=[4.04, 23.30]; $\chi^2(1) = 12.43$, $p < .001$), such that participants were overall more accurate on Mid-Decade items (M=49%, SD=34%) than Decade Transition items (M=17%, SD=30%).

Finally, we test the prediction that Productive Counters can generate successors for numbers both within and beyond their Initial Highest Count, while Non-Productive Counters are only able to generate successors within their Initial Highest Count. This regression model predicted trial-level accuracy from Productivity, Item Range (Within/Beyond IHC), a Productivity by Item Range interaction, as well as Initial Highest Count and age. Here, we found a significant main effect of Productivity (β = .71, OR=2.04, 95% CI = [1.00, 4.16]; $\chi^2(1) = 3.19$, $p = .0495$) and Initial Highest Count (β = .88, OR=2.40, 95% CI = [1.59, 3.61]; $\chi^2(1) = 17.56$, $p < .001$), but no significant effect of Item Range ($\chi^2(1) < 1$, $p = .99$). Planned contrasts indicate that performance on numbers within children's Initial Highest Count was similar for Productive Counters (IHC<99) ($M = 60\%$) and Non-Productive Counters ($M = 62\%$; $p = .86$ by *t*-test) whereas accuracy for numbers beyond their Initial Highest count was significantly greater among Productive Counters ($M = 55\%$, $SD = 30\%$) than among Non-Productive Counters ($M = 30\%$, $SD = 29\%$; $t(88) = -3.96$, $p < .001$). However, this Item Range by Productivity interaction did not meet the threshold for significance (β = .94, OR=2.57, 95% CI = [.96, 6.87], $\chi^2(1) = 3.54$, $p = .06$).

D3: Relationship of Productivity to Infinity knowledge

Here we take the same regression approach as described in the main text, replacing the variable of Productivity (Non-Productive / Productive (IHC < 99)) with the stricter version. In predicting children's Successor Knowledge (Table S5) and Endless Knowledge (Table S6), none of the three predictors explained a significant proportion of additional variance compared to the base model.

Finally, we constructed models predicting children's status as Full Infinity Knowers (Table S7). Productivity Group explained significant additional variance relative to the base model ($\chi^2(1) = 4.67$, $p = .031$), though other measures of counting ability (Initial Highest Count, Next Number accuracy) did not explain additional variance when controlling for age. This final model thus included only Productivity Group and Age as predictors, and estimated that Productive counters were more likely than Non-Productive Counters to have Endless Knowledge ($\beta = 1.70$, $p = .04$, OR = 5.47, 95% CI: 1.17–31.80), while Age was not a significant predictor ($p = 0.83$). The effect of Productivity Group remained when controlling for Initial Highest Count, although these more complex models did not explain any additional variance.

Table S5

Regression models for predicting Successor knowledge on the Infinity Interview (Participants IHC <99, N=90)

Models	Coefficient Estimates (ß)				Summary statistics		
	Age	IHC	Next Number accuracy	Productivity Group	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model							
Age	0.296				-58.75	121.51	0.027
Initial Models							
Age + IHC	0.598 *	-0.520			-57.05	120.10	0.077
Age + Next Number accuracy	0.358		-0.155		-58.55	123.10	0.033
Age + Productivity Group	0.056			0.885	-57.39	120.77	0.067

Notes. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.

[a] Each initial model was compared against the base model.

* $p < 0.05$

Table S6

Regression models for predicting Endless Infinity knowledge on the Infinity Interview (Participants IHC <99, N=90)

Models	Coefficient Estimates (ß)				Summary statistics		
	Age	IHC	Next Number accuracy	Productivity Group	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model							
Age	0.451				-42.22	88.44	0.049
Initial Models							
Age + IHC	0.261	0.349			-41.63	89.26	0.070
Age + Next Number accuracy	0.287		0.466		-41.04	88.07	0.090
Age + Productivity Group	0.133			1.171	-40.66	87.32	0.103

Notes. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.

[a] Each initial model was compared against the base model.

Table S7

Regression models for predicting Full knowledge on the Infinity Interview (Participants IHC <99, N=90)

Models	Coefficient Estimates (ß)				Summary statistics		
	Age	IHC	Next Number accuracy	Productivity Group	Loglikelihood [a]	AIC	$R^2_{Nagelkerke}$
Base Model							
Age	0.527				-33.91	71.81	0.058
Initial Models							
Age + IHC	0.481	0.084			-33.88	73.76	0.059
Age + Next Number accuracy	0.417		0.307		-33.51	73.03	0.073
Age + Productivity Group	0.082			1.700 *	-31.57*	69.14	0.148
Final Models							
Age + Productivity Group + IHC	0.132	-0.133		1.772 *	-31.53	71.05	0.149
Age + Productivity Group * IHC	0.143	-0.066		1.777 *	-31.41	72.82	0.154

Notes. Coefficients were compared against 0 using *t*-tests. Model comparisons done using Likelihood Ratio Tests.

[a] Each initial model was compared against the base model.

* $p < 0.05$, ** $p < 0.01$

Open Practices Disclosure

Manuscript Title: Counting to Infinity: Does learning the syntax of the count list predict knowledge that numbers are infinite?
Corresponding Author:

Articles accepted to *Cognitive Science* after November 2018 are eligible to earn badges that recognize open scientific practices: publicly available data or material. Please read more about the badges in our author guidelines, and you can also find information on the Open Science Framework wiki.

[x] Please check this box if you are interested in participating.

To apply for one or more badges acknowledging open practices, please check the box(es) corresponding to the desired badge(s) below and provide the information requested in the relevant sections. To qualify for a badge, you must provide a URL, doi, or other permanent path for accessing the specified information in a public, open-access repository. **Qualifying public, open-access repositories are committed to preserving data, materials, and/or registered analysis plans and keeping them publicly accessible via the web in perpetuity.** Examples include the Open Science Framework (OSF) and the various Dataverse networks. Hundreds of other qualifying data/materials repositories are listed at http://re3data.org/. **Personal websites and most departmental websites do not qualify as repositories.**

Authors who wish to publicly post third-party material in their data or materials must have the proper authority or permission agreement in order to do so.

There are circumstances in which it is not possible or advisable to share any or all data or materials publicly. For example, there are cases in which sharing participants' data could violate confidentiality. If you would like your article to include an explanation of such circumstances and/or provide links to any data or materials you have made available—even if not under conditions eligible to earn a badge—you may write an alternative note that will be published in a note in the article**.** Please check this box if you would like your article to include an alternative note and provide the text of the note below:

[] Alternative note:

[x] Open Data Badge

1. Provide the URL, doi, or other **permanent path** for accessing the data in a **public, open-access repository**: https://osf.io/z6ky3/

[x] Confirm that there is sufficient information for an independent researcher to reproduce **all of the reported results**, including codebook if relevant.

[x] Open Materials Badge

1. Provide the URL, doi, or other **permanent path** for accessing the materials in a **public, open-access repository**: https://osf.io/z6ky3/

[x] Confirm that there is sufficient information for an independent researcher to reproduce **all of the reported methodology**.

By signing below, authors affirm that the above information is accurate and complete, that any third-party material has been reproduced or otherwise made available only with the permission of the original author or copyright holder, and that publicly posted data do not contain information that would allow individuals to be identified without consent.

Signature: Junyi Chu

Name: Junyi Chu

Date: February 26, 2020