Recursive Statistical Learning in Adults, Children, and Macaques

Elijah Tramm (etramm@wisc.edu)[1] & Stephen Ferrigno[1]
[1]Department of Psychology, University of Wisconsin–Madison

Abstract

Humans readily learn language, mathematics, music, and a host of other sequentially structured conceptual systems beginning in early childhood. By contrast, even with extensive training, other species learn only the basics of many of these domains. One proposed explanation for this difference is that only humans are capable of using recursive operations to learn complex sequential concepts. Here, we build on recent work on recursive cognition to test the evolutionary and developmental origins of this ability. We tested adults, children, and rhesus macaques on a binary sequence learning task that can be learned using statistical learning recursively. Adults readily applied this recursive statistical learning strategy, but children did not, and macaques showed mixed evidence of doing so. We consider the possible reasons for the results of the children and macaques, as well as their possible implications for theories of cognition.

Keywords: recursion; statistical learning; comparative cognition; cognitive development

Introduction

Language, mathematics, logic, music, and complex theory of mind are a few of the hallmarks of uniquely human intelligence. Other species display only rudimentary capabilities in these domains, whereas humans tend to learn them beginning in early childhood. What is it about human cognition that enables these singularly human abilities?

Some have taken the stance that human cognition is qualitatively different from nonhuman cognition. On this view, humans have evolved representational capacities or computational processes that allow us to go beyond the cognitive abilities of other species. One proposal for such an ability that has received extensive attention is recursion, a computational process in which an operation is applied to its own output (Fitch, 2010).

Hauser et al. (2002) were among the first to popularize recursion as a uniquely human adaptation that led to new cognitive abilities. According to their proposal, recursion is the computational mechanism that enables the creative construction of infinitely many meaningful structures. They suggest that recursion is responsible for language, acting as the compositional engine behind syntax. They also suggest other roles for recursion in uniquely human cognition, such as in learning the principles behind number.

Many subsequent studies attempted to evaluate this theory, often by examining how human and nonhuman animals comprehend and produce center-embedded structures (Ferrigno, 2022). Non-human primates have been shown to represent a variety of sequential structures and relations, including adjacent and non-adjacent dependencies (see Santolin & Saffran, 2014 and Wilson et al., 2020 for reviews), abstract patterns (Wang et al., 2015), and hierarchical dependencies (see Ferrigno, 2022 for review). However, research to date has had difficulty in determining whether recursive computations underlie any of these sequence processing abilities. Although this research has produced important findings, center-embedded sequences are not the only avenue to investigate the recursive capacities of humans and other animals.

One recently developed alternative focuses on recursion as an operation in a Language of Thought (LoT), a sort of mental programming language that instantiates the computations involved in cognition (Dehaene et al., 2022). In LoT models, recursion plays a similar role to recursion in computer programming, allowing operations to call themselves or be applied to their own output. Generally, recursion has not been the focus of these models. It is usually included as an assumption about the LoT (e.g., Yildirim & Jacobs, 2015). However, recent work has begun investigating the recursive aspect of these models.

Planton et al. (2021) chose to address the recursive aspect of LoT models more specifically by investigating how adults learn binary sequences. They created binary sequences that differed in complexity as measured by the minimum description length in a simple LoT with the operations stay and switch and the option to recursively call those operations. This LoT complexity could be compared to other complexity measures such as surprise (based on transitional probability), chunk complexity (Mathy & Feldman, 2012), and entropy, to see which best accounted for adult performance in a sequence oddball detection paradigm. They found that LoT complexity was most strongly correlated with performance and subjective complexity ratings.

However, Planton et al. (2021) note that it is unlikely that humans could effectively search the hypothesis space of potential generative programs to find the shortest possible one. Furthermore, they found that participants showed a reticence to divide chunks, which led to worse fit when the LoT model required the division of chunks. Based on these factors, they proposed a more plausible process in which participants first combine runs of identical items into chunks, and only then use the LoT to further compress the sequence. They found that this LoT and chunking measure accounted for performance even better than the LoT measure.

In Planton et al.'s study, most of the sequences required only a little more than one level of embedding (2021). This makes it difficult to determine whether the mechanism participants used is truly recursive. At low levels of embedding, non-recursive strategies can often achieve the same results as recursive ones; at greater levels of embedding, these strategies break down. Thus, a critical test of whether a given process is truly recursive is whether it generalizes to higher levels of embedding (Martins, 2012).

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

A similar line of work used a different binary sequence with the potential to overcome this limitation. These experiments used a unique form of binary sequences generated by L-grammars (Lindenmayer, 1968), which differ from grammars on the Chomsky hierarchy in that all rewrite rules are applied simultaneously, and terminal symbols are defined by sequence context (Krivochen et al., 2018). This grammar, known as 'Fib', produces sequences recursively by concatenating the two previous generations of the sequence[1], similar to how the Fibonacci sequence is generated. This results in sequences with the special properties of self-similarity and aperiodicity. Self-similarity refers to the fact that these sequences are made up of chunks of natural constituents that share similar features. Aperiodicity refers to the fact that the sequences do not repeat themselves, i.e., there is not a linear function that can predict upcoming points with certainty. These sequences are hierarchically nested, with each generation being made up of the two previous generations, which in turn are made up of the two previous generations, and so on. Geambaşu et al. (2016; 2020) and Vender et al. (2020) showed that adults and children as young as ten-years-old can learn these sequences across modalities.

Schmid et al. (2023) proposed a recursive mechanism that could construct mental representations of these sequences via statistical learning. According to their theory, participants use statistical learning recursively to build a compressed representation of the sequence. As participants are exposed to the sequence (displayed as red and blue circles), they discover basic transitional probabilities (TPs; e.g., p(red|blue) = 1.0). In Fib sequences, TPs can be 0.38, 0.62, or 1.0. The deterministic (1.0) transitions can be used to group items into chunks, forming a single, unitary item in the sequence (e.g., [blue] and [red] → [blue, red], see Figure 1). While the probabilistic transitions could, in principle, also be used for chunking, this would interfere with the recursive statistical learning process because the probabilistic transitions change based on the size of the chunk created. Therefore, if participants do appear to represent chunks it is unlikely they also track the changing probabilistic transitions.

To further reduce uncertainty and compress the sequence, statistical learning can be applied recursively to the chunked representation of the sequence created by the initial application of statistical learning. In this newly chunked representation, the possible TPs are the same (i.e., they can still only be .38, .62, or 1.0). However, the probability of each transition changes, and some transitions are newly deterministic; the transition from the first red to the first blue was non-deterministic in the base sequence but is now deterministic. The TPs of this new chunked sequence are independent from those of the base sequence. Re-applying statistical learning thus reveals new deterministic transitions, making the sequence more predictable. This recursive statistical learning process can, in theory, be applied ad infinitum, each time revealing new deterministic transitions (Figure 1).

[1] Alternatively, sequences can be generated using the rewrite rules 0 → 1 and 1 → 01.

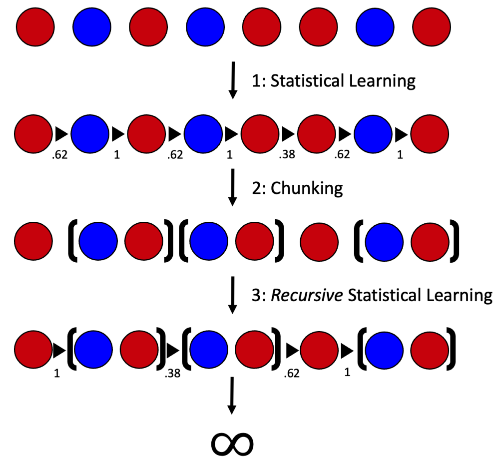

Figure 1: Diagram of the recursive statistical learning process. Starting with the base sequence, the learner uses statistical learning to discover TPs between items in the sequence. These TPs are then used to chunk the sequence. In the next step, which makes the process recursive, statistical learning is re-applied to the chunked sequence created by the previous application of statistical learning. Doing so reveals that some TPs have changed, e.g., the TP from red to blue was .62, but the TP from red to [blue, red] is 1.0.

This paradigm offers two advantages over that of Planton et al. (2021). First, it can, in theory, test an infinite depth of recursive processing, allowing for a more precise delineation of recursive capacity. Second, rather than relying on the LoT framework, it is based in the more broadly accepted mechanism of statistical learning. This allows us to test if participants have the capacity to use recursive statistical learning, as well as measure the number of possible embeddings that can be represented.

To determine whether adult participants used this recursive statistical learning strategy, and if so to what extent, Schmid et al. (2023) employed a serial reaction time task. They used generation 12 of the sequence, which is 233 items long, and repeated it for five blocks. Each trial, an item from the sequence appeared on a screen, and participants had to press a corresponding key as quickly as possible. If participants were using recursive statistical learning, they should have been able to predict upcoming items that followed deterministic transitions, leading to faster reaction times.

In line with this hypothesis, Schmid et al. (2023) found that participants showed faster reaction times for deterministic transitions revealed by three applications of the recursive statistical learning process. Schmid et al.'s (2023) paradigm is an exciting extension of Planton et al.'s (2021) study both because it can test greater depths of recursive processing, and because it is not based on the LoT framework, which is not universally accepted.

This work represents a promising avenue for investigating the presence and extent of recursive computations in human and nonhuman cognition. Dehaene et al. (2022) claim that work by Amalric et al. (2017) and Planton et al. (2021), along with other studies, suggests that humans are endowed with a unique ability to create mental programs using recursive operations in a LoT. It remains an open question whether nonhuman animals are capable of using recursive operations.

Here, we used a modified version of Schmid et al.'s recursive statistical learning task to compare the capacity for recursive statistical learning in adults, children, and rhesus macaques (*Macaca mulatta*).

Methods

Participants

Participants were 40 adults, 30 children, and four rhesus macaques. Participant sample size was not preregistered.

Adult participants were recruited from the University of Wisconsin-Madison undergraduate psychology participant pool. Children aged three- to seven-years old (M_{age}: five years, nine months, σ_{age}: two years, 11 months) were recruited via the Cognitive Origins Lab database. Participants or their legal guardians gave informed written consent to participate in the study. Macaques (all female) aged eight- to 10-years old (M_{age}: nine years, eight months, σ_{age}: 10 months) were pair-housed, and fed monkey chow, fruits, vegetables, and water ad libitum. All procedures were approved by the University of Wisconsin-Madison Institutional Animal Care and Use Committee. Macaques completed testing in their home enclosures with a touchscreen computer and pellet dispenser attached to a mesh grating.

Materials and Procedure

The test sequence for this experiment was a 233-item long generation of the Fib grammar. Participants saw a gray screen with two black circle outlines, one centered on each half of the screen (Figure 2). A trial consisted of a colored circle appearing within one of these outlines. The color of the circle was determined by the binary sequence. Blue circles always appeared in the left outline, and red circles always appeared in the right outline. Because only one circle appeared on screen in a given trial, it was not possible to make errors in the task. While this precluded accuracy data, we chose this design to make the task more conducive to learning and performance in macaques. To complete a trial, the participant had to tap the item, which caused it to disappear. If the item was not tapped within .75 seconds, it began to decrease in opacity, completely disappearing after five seconds. If this occurred, a buzzer played and the screen would become red for five seconds. Following a variable ratio of 18 to 22 trials, a chime would play and the screen would become green. This would mark a five-second break, after which trials resumed.

Adults completed five blocks of the 233-item long sequence, as in Schmid et al. (2023). We chose to replicate the general structure of Schmid et al. to ensure that the modifications we made to the task did not prevent participants from learning. Because children became disinterested in the task more quickly, they completed only one block. Due to the smaller sample size of macaques, we had them complete at least 30 sessions of five blocks.

Adults were told that they would be completing a reaction time task, and that their goal was to tap the items on the screen as quickly as possible once they appeared. Adults completed testing in a quiet room on an upright touchscreen tablet. Testing took approximately 25 minutes.

Children were told that they were going to play a game similar to whack-a-mole, where they had to tap the figures on the screen as quickly as possible to pass a level (i.e., reach a reward pause). During the reward pause, they were allowed to choose a sticker to add to their collection. Once they completed the task, they were allowed to choose a stuffed animal and $10 as a prize along with their stickers. The duration of the test depended on the child's performance, with testing generally taking between 25 and 40 minutes.

All macaques were experimentally naïve, so they had to be trained to perform the reaction time task. We began with a habituation task teaching the macaques to press a single icon in the center of the screen to receive a reward. Once they learned this, we gave them a modified version of the reaction time task in which, rather than every 18 to 22 items, a reward was given every three to five items. Once macaques completed one testing session of this, the reinforcement ratio was increased to every five to nine items, then to every 10 to 16 items. Once they completed this training, they moved on to the final task.

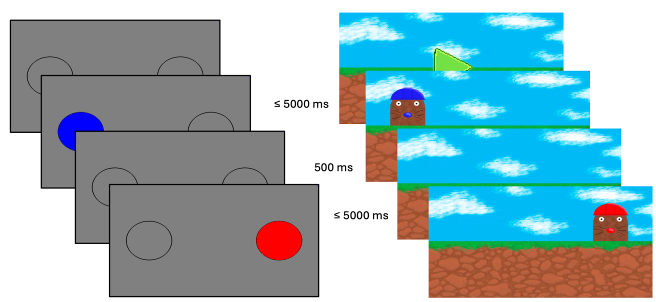

Figure 2: Diagram of trial structure. One of the stimuli corresponding to a binary item from the grammar appeared on screen, remaining for up to five seconds. When tapped, it disappeared, beginning a 500 millisecond intertrial interval.

Analysis

We began by restricting our analysis to trials in which reaction times were within two standard deviations of the mean for each subject in order to remove outliers.

Our interest was whether participants were able to use recursive statistical learning to parse the sequence. Recursive statistical learning reveals specific TPs with each application. If participants were using recursive statistical learning, they should be able to use the TPs to predict upcoming items, in turn allowing them to react more quickly to them.

The primary predictor of interest for our analyses was determinacy status (i.e., whether a given transition was deterministic or non-deterministic). Each application of the recursive statistical learning process results in a representation made up of new, larger chunks. When recursive statistical learning is then applied to this new, chunked representation, some transitions are revealed to be newly deterministic, while others remain probabilistic. We compare reaction times for the deterministic transitions to the new chunks to non-deterministic transitions to the new chunks. For example, in Figure 1, only transitions to a [blue, red] chunk are analyzed. Some of these are deterministic, while others are non-deterministic. We use these transitions to compare reaction times for the same items, in different contexts: deterministic vs. non-deterministic transitions. By comparing reaction times for these particular transitions, we are able to examine how the structural context of the chunk influences reaction time while controlling for the effects of previous applications of statistical learning.

For adults, we fit linear mixed-effects models predicting reaction time using determinacy status, block, and their interaction. Determinacy status was a dichotomous variable (deterministic or non-deterministic) and block was a mean centered continuous variable. We used the maximal model, which included a by-subject random intercept, and by-subject random slopes for determinacy status, and block, with their interaction being dropped to allow for convergence.

For children, we fit linear mixed-effects models predicting reaction time using determinacy status, trial number, and their interaction. Trial number was mean centered. We also included a by-subject random intercept and by subject random slopes for determinacy status, trial number, and their interaction.

We fit similar linear models for each macaque, predicting reaction time using determinacy status and session number, and their interaction. Determinacy status and block were coded the same as in the models for adults and children. Session number was a mean centered continuous variable. Because we analyzed data for macaques individually, we did not include a by-subject random effects.

Results

Adults

We began by fitting a model to determine whether adults were using classical statistical learning on the task. We found that they were, with reaction times for deterministic transitions being faster than reaction times for non-deterministic transitions ($b = 129.71$, $SE = 8.5$, $p < .001$). This confirmed that our modified task did not interfere with adults' ability to use statistical learning or our ability to detect it.

Next, we looked to see whether adults applied statistical learning to the chunks created by the application of basic statistical learning, i.e., whether they were using recursive statistical learning. Adults reacted more quickly to transitions that were revealed as deterministic after one application of recursive statistical learning than to transitions that were non-deterministic ($b = 88.11$, $SE = 11.4$, $p < .001$), suggesting that they were applying the recursive statistical learning strategy.

We then ran models to determine whether they continued the process of applying statistical learning to the chunks generated by the previous application of statistical learning. In line with the recursive statistical learning strategy, adults also reacted more quickly to deterministic transitions revealed by two ($b = 32.49$, $SE = 7.08$, $p < .001$) and three ($b = 20.79$, $SE = 5.85$, $p < .01$) applications of the strategy, but not four ($b = -7.32$, $p = .076$) (Figure 3). These results show clear evidence of recursive statistical learning with three applications, replicating the results of Schmid et al. (2023). The decreasing magnitude of the difference in reaction time between deterministic and non-deterministic transitions is seen in the individual data, suggesting the benefit of this recursive process decreases with each application.

Each model also included trial block and the interaction between determinacy status and trial block as predictors. The effect of block was significant for all models, with reaction times decreasing overall as participants gained experience with the task (all *ps* $\leq .001$). The interaction between trial block and determinacy status was also significant for applications 1 ($b = 17.26$, $SE = 3.44$), 2 ($b = 8.67$, $SE = 3.19$), and 3 ($b = 10.51$, $SE = 3.12$) (all *ps* $\leq .001$), but not application 4 ($b = 3.95$, $p = .168$), suggesting that participants improved in their ability to use the deterministic transitions revealed by each application of recursive statistical learning they discovered.

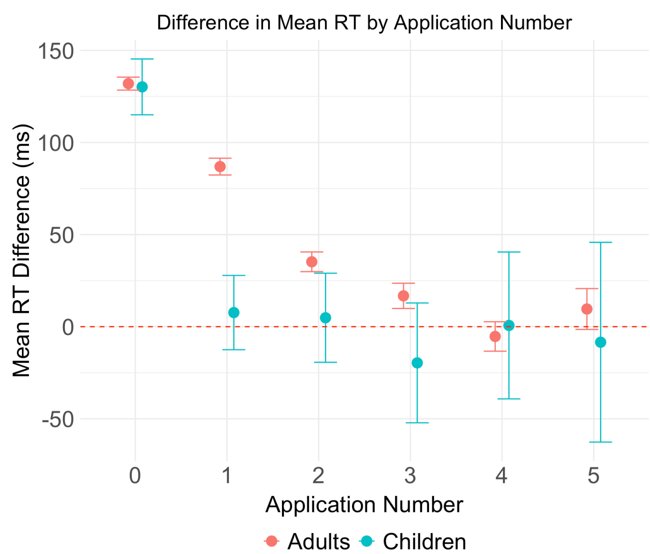

Figure 3: Mean reaction time difference for deterministic and non-deterministic transitions as revealed by each application of statistical learning for human participants. Error bars represent 95% confidence intervals.

Children

Following the same steps as the data analyses of adults, we began by determining whether children were using classical statistical learning on the task. Overall, children reacted more quickly to deterministic transitions than to non-deterministic transitions ($b = 128.37$, $SE = 11.19$, $p < .001$) (Figure 3). This confirmed that children used statistical learning to predict upcoming items in the task. There was also a significant effect of trial and the interaction between trial and determinacy status, such that children became faster with experience.

We then fit models to examine whether children used recursive statistical learning to further learn the sequence. We found no significant effect of determinacy status for applications one (b = 9.51), two (b = 5.88) or three (b = -19.98) (all *ps* ≥ .117). This suggests that children were not recursively applying statistical learning to the outputs of previous applications. For both of these models, the effect of trial number was significant, but the interaction between trial number and determinacy status was not significant. Thus, children did not appear to learn to use the recursive statistical learning process with experience.

We also fit models which included children's age (as a continuous mean centered variable) and the interaction between age and determinacy status as variables to determine whether children developed the ability to use recursive statistical learning with age[2]. There was no significant effect of age or the interaction between age and determinacy status on reaction time.

Although children did not show evidence of recursive statistical learning, it is possible that the recursive strategy couldn't be learned in the smaller number of trials children experienced relative to adults. To determine whether the strategy could be learned and applied in this span, we analyzed the results from adult participants but restricted our analysis to only the first 233 trials, the same number children completed. We found that there was a significant effect of determinacy status on reaction time for deterministic transitions revealed by 1 (b = 40.15) and 2 (b = 12.38) (*ps* ≤ .005) applications of the recursive statistical learning process. This suggests that children's failure to use the recursive strategy was not due to an inherent impossibility in doing so.

Macaques

For macaques, we followed the same analysis procedure we used for adults. However, we fit models for each macaque individually, as we were interested in the ability of any macaque to use recursive statistical learning, not the prevalence of this ability among macaques.

We first fit models to determine whether macaques were using basic statistical learning on the task. We found that three of the four macaques reacted more quickly to deterministic transitions than non-deterministic transitions (Macaque (M) 2: b = 54.61, SE = 2.94; M3: b = 176.03, SE = 3.58; M4: b =125.50, SE = 5.33; *ps* ≤ .001). This confirmed that some macaques were using statistical learning on the task. We stopped analyzing the data from the macaque that did not use statistical learning, as it couldn't have used recursive statistical learning.

We then fit models to determine whether macaques applied statistical learning recursively. All three macaques that used statistical learning also reacted more quickly to deterministic transitions revealed by one application of recursive statistical learning (M2: b = 50.98, SE = 3.71; M3: b = 48.76, SE = 4.73; M4: b = 26.40, SE = 8.02; *ps* ≤ .001).

Next, we analyzed whether the macaques were able to apply this recursive statistical learning process multiple times, as human adults did. Macaques do not appear to have applied the recursive strategy multiple times, with reaction times for deterministic transitions revealed by two or three applications of the process being similar to or even slower than those for non-deterministic transitions (Figure 4).

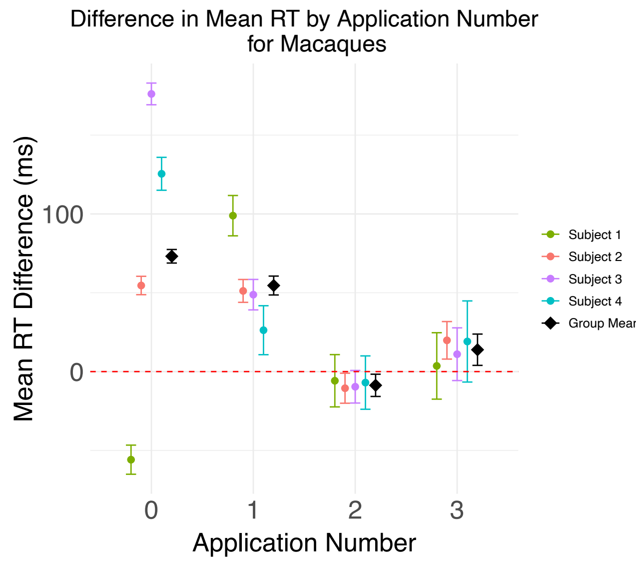

Figure 4: Mean reaction time difference for deterministic and non-deterministic transitions as revealed by each application of statistical learning in macaques. Error bars represent 95% confidence intervals.

Because macaques completed multiple sessions of the task, we also examined the effect of session and the interaction between determinacy status and session on reaction time in each of these models. The effect of session was significant in all models, suggesting that macaques became faster with experience in the task. The interaction between session and determinacy status was significant for the transitions revealed by basic statistical learning ($p < .001$), but not for any transitions revealed by recursive statistical learning.

Discussion

In this study, we sought to investigate whether adults, children, and nonhuman primates could implement a recursive statistical learning computation. To do this, we implemented a modified version of a paradigm developed by

[2] Data from two children had to be excluded due to their ages not being recorded properly.

Schmid et al. (2023) to test these three populations in the same way.

Replicating the results of Schmid et al. (2023), we found strong evidence of recursive statistical learning in adults. Adults reacted more quickly to transitions that were revealed as deterministic after three applications of the recursive statistical learning process, suggesting that they were sensitive to the TPs. Aside from replicating the results of Schmid et al., this confirms that the modifications we made to the task did not interfere with learning.

We found no evidence of recursive statistical learning in children. Although the current sample size limits the conclusions we can draw, if these results hold, they would be surprising under any theory which uses recursion to explain how children learn many uniquely human forms of cognition (e.g., Corballis, 2007). There are several ways to interpret our results, though we note that none of them can be fully endorsed based on the current evidence.

The first possible interpretation is that our results demonstrate a genuine inability of children to implement recursive operations. If this is the case, theories utilizing recursion to explain children's acquisition of complex concepts would need to be revised to explain how children can acquire these concepts without recursive computations. Given that adults displayed clear recursive ability, a further question arising from this interpretation is when and how the capacity for recursion develops.

The second possibility is that mastering recursive computations requires more experience. The adult level of performance could reflect years of cumulative experience with recursive computations in language, mathematics, and other domains.

The third possible interpretation is that our results are due to a domain specificity of recursion early in life. For example, perhaps recursion is at first only available for linguistic computations but becomes available for other computations over development.

The fourth possible interpretation is that children's failure was the result of interference from other confounds. Factors such as the amount of exposure, motivation, or poor motor coordination could prevent children from enacting predictions that they might have made. The specific nature of the sequence might also have made learning difficult for children. Prior work on statistical learning has shown that children struggle to segment sequential input into chunks of variable length (Lew-Williams & Saffran, 2012). Since chunk length varies when recursive statistical learning is used, this may have interfered with learning.

Of the four macaques we tested, three showed reaction time patterns consistent with recursive statistical learning. Each of these three macaques reacted more quickly to transitions that were revealed as deterministic after re-applying statistical learning to the chunks created by an initial application of statistical learning. Macaques' performance thus appears slightly better than that of human children, but worse than that of human adults.

This intermediate performance leads to some difficulty in interpreting the results. An important test of whether a given process is recursive is whether it generalizes to higher levels of embedding, as non-recursive strategies begin to break down at higher levels (Martins, 2012; Planton et al. 2021). Although macaques showed signs of applying the recursive statistical learning process, we cannot easily determine whether they were using the recursive strategy. The results of a single application of the recursive statistical learning strategy can also be accomplished with non-recursive strategies, as can any recursive operation that is not applied infinitely. However, with multiple applications, the types of strategies that could account for the data reduces substantially. Non-recursive models, such as recurrent neural networks, can also produce the output of a single application of recursive statistical learning in this task (Lakretz & Dehaene, 2021). This does not mean we should reject the possibility that they were. First, similar to children, macaques have little experience with recursive structures, putting them at a disadvantage. Second, the same limitations are seen in examples commonly cited as evidence of recursive computations in human cognition as well. Despite linguistic embedding being limited to a depth of three in adults (Karlsson, 2007), the process is often still taken to be recursive, with the limitations being attributed to working memory rather than the recursive process itself (Miller & Chomsky, 1968). The present situation can be interpreted similarly. Macaques may be able to use recursive computations, but only to a depth of one application due to working memory limitations.

Such a result would be expected if, as one recently proposed theory suggests, human cognitive uniqueness is the result of increased information processing capacity rather than novel computational abilities (Cantlon & Piantadosi, 2024). An important focus for future research will be to distinguish between limitations due to the use of non-recursive processes and limitations due to working memory.

Conclusion

Our results highlight the continued need for research on the evolutionary and developmental origins of recursive computations in cognition. Adults successfully implemented recursive statistical learning, but only three times on average, suggesting that this ability is rather limited even with simple stimuli. Children did not seem to use recursive statistical learning, although this result is not decisive due to the differences in experimental conditions. Given the importance of recursion to some theories of conceptual development, further examination of this failure is critical. Macaques showed signs of one application of recursive statistical learning, although their performance may be explained in other ways. If these results do represent success on the part of the macaques, it would challenge the assumption that recursion is unique to humans. Altogether, these findings suggest that further research on recursion in cognition is needed to ground theories and models in empirical data.

Acknowledgements

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health under Award Number R01HD115600 (to SF). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would also like to acknowledge support from the American Society of Primatologists (to ET). We would like to thank Maimoona Abdaal and the Cognitive Origins Lab for research administration and help collecting data.

References

Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., & Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers. PLOS Computational Biology, 13(1), e1005273. https://doi.org/10.1371/journal.pcbi.1005273

Cantlon, J. F., & Piantadosi, S. T. (2024). Uniquely human intelligence arose from expanded information capacity. Nature Reviews Psychology. https://doi.org/10.1038/s44159-024-00283-3

Corballis, M. C. (2014). The recursive mind: The origins of human language, thought, and civilization. Princeton University Press.

Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022). Symbols and mental programs: A hypothesis about human singularity. Trends in Cognitive Sciences, 26(9), 751–766. https://doi.org/10.1016/j.tics.2022.06.010

Ferrigno, S., Cheyette, S. J., Piantadosi, S. T., & Cantlon, J. F. (2020). Recursive sequence generation in monkeys, children, U.S. adults, and native Amazonians. Science Advances, 6(26), eaaz1002. https://doi.org/10.1126/sciadv.aaz1002

Ferrigno, S. (2022). Sequencing, Artificial Grammar, and Recursion in Primates. In B. L. Schwartz & M. J. Beran (Eds.), Primate Cognitive Studies (1st ed., pp. 260–290). Cambridge University Press. https://doi.org/10.1017/9781108955836.011

Fitch, W. T. (2010). Three meanings of "recursion": Key distinctions for biolinguistics. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), The Evolution of Human Language (1st ed., pp. 73–90). Cambridge University Press. https://doi.org/10.1017/CBO9780511817755.005

Geambaşu, A., Ravignani, A., & Levelt, C. C. (2016). Preliminary experiments on human sensitivity to rhythmic structure in a grammar with recursive self-similarity. Frontiers in Neuroscience, 10, 281.

Geambaşu, A., Toron, L., Ravignani, A., & Levelt, C. C. (2020). Rhythmic recursion? Human sensitivity to a Lindenmayer grammar with self-similar structure in a musical task. Music & Science, 3, 2059204320946615.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? Science, 298(5598), 1569–1579. https://doi.org/10.1126/science.298.5598.1569

Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. Journal of Linguistics, 43(2), 365–392. https://doi.org/10.1017/S0022226707004616

Krivochen, D., Phillips, B., & Saddy, D. (2018). Classifying points in Lindenmayer systems: transition probabilities and structure reconstruction. Ms. University of Reading.

Lakretz, Y., & Dehaene, S. (2021). Recursive processing of nested structures in monkeys? Two alternative accounts. Science Advances, 6(26), 1002.

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. Cognition, 122(2), 241–246. https://doi.org/10.1016/j.cognition.2011.10.007

Lindenmayer, A. (1968). Mathematical models for cellular interactions in development I. Filaments with one-sided inputs. Journal of theoretical biology, 18(3), 280-299.

Martins, M. D. (2012). Distinctive signatures of recursion. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1598), 2055–2064. https://doi.org/10.1098/rstb.2012.0097

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. Cognition, 122(3), 346–362. https://doi.org/10.1016/j.cognition.2011.11.003

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users.

Planton, S., Van Kerkoerle, T., Abbih, L., Maheu, M., Meyniel, F., Sigman, M., Wang, L., Figueira, S., Romano, S., & Dehaene, S. (2021). A theory of memory for binary sequences: Evidence for a mental compression algorithm in humans. PLOS Computational Biology, 17(1), e1008598. https://doi.org/10.1371/journal.pcbi.1008598

Santolin, C., & Saffran, J. R. (2018). Constraints on Statistical Learning Across Species. Trends in Cognitive Sciences, 22(1), 52–63. https://doi.org/10.1016/j.tics.2017.10.003

Schmid, S., Saddy, D., & Franck, J. (2023). Finding Hierarchical Structure in Binary Sequences: Evidence from Lindenmayer Grammar Learning. Cognitive Science, 47(1), e13242. https://doi.org/10.1111/cogs.13242

Vender, M., Krivochen, D. G., Compostella, A., Phillips, B., Delfitto, D., & Saddy, D. (2020). Disentangling sequential from hierarchical learning in Artificial Grammar Learning: Evidence from a modified Simon Task. PLOS ONE, 15(5), e0232687. https://doi.org/10.1371/journal.pone.0232687

Wang, L., Uhrig, L., Jarraya, B., & Dehaene, S. (2015). Representation of Numerical and Sequential Patterns in Macaque and Human Brains. Current Biology, 25(15), 1966–1974. https://doi.org/10.1016/j.cub.2015.06.035

Wilson, B., Spierings, M., Ravignani, A., Mueller, J. L., Mintz, T. H., Wijnen, F., Van Der Kant, A., Smith, K., & Rey, A. (2020). Non-adjacent Dependency Learning in Humans and Other Animals. Topics in Cognitive Science, 12(3), 843–858. https://doi.org/10.1111/tops.12381

Yildirim, I., & Jacobs, R. A. (2015). Learning multisensory representations for auditory-visual transfer of sequence category knowledge: A probabilistic language of thought approach. Psychonomic Bulletin & Review, 22(3), 673–686. https://doi.org/10.3758/s13423-014-0734-y