A Biologically Plausible Model of Mental Multiplication

Isaac Joffe (ijoffe@uwaterloo.ca)[1,2] **& Chris Eliasmith**[1,2,3,4]
[1]Centre for Theoretical Neuroscience, University of Waterloo
[2]David R. Cheriton School of Computer Science, University of Waterloo
[3]Department of Philosophy, University of Waterloo
[4]Department of Systems Design Engineering, University of Waterloo

Abstract

Mental multiplication is an abstract and advanced cognitive skill that requires applying systematic algorithms to structured numerical data, but it remains underexplored in neural models. In this paper, we propose a psychologically plausible model of how anatomical circuits in the human brain may perform mental multiplication of small natural numbers and implement it in biologically plausible spiking neurons. Our model uses similar strategies to those used by humans, specifically repeated addition and rule use; matches various human performance levels, including that of both children and adults; and qualitatively replicates multiple human performance patterns, such as problem-size and outlier effects. Our novel model sheds further light on how structured cognitive processes, like the rule-based algorithms underlying mathematical reasoning, may be implemented in the substrate of spiking neurons in the human brain.

Keywords: Mental multiplication; mental arithmetic; cognitive modelling; spiking neural networks; semantic pointer architecture; neural engineering framework

Introduction

Mental multiplication, despite its ostensible simplicity and ubiquity, is an abstract and advanced cognitive skill. Although some non-human animals have approximate number sense and can perform simple arithmetic tasks (Boysen & Berntson, 1989; Brannon, 2005; Cantlon & Brannon, 2007), there is no evidence of them performing multiplication. Additionally, contemporary artificial intelligence (AI) systems (e.g., large language models) struggle with mathematical reasoning (Hendrycks et al., 2021; Yuan et al., 2023) and representing systematic numerical relationships (Ouellette et al., 2024; Yehudai et al., 2024). Furthermore, even humans only learn multiplication after years of schooling (Schön et al., 2012; Steel & Funnell, 2001). Put simply, the ability to perform mental multiplication separates mature humans from animals, AI, and young children.

In this paper, we propose a novel, psychologically and biologically plausible spiking neural model of simple mental multiplication. Although cognitive models of mental arithmetic exist (e.g., Griffiths & Kalish, 2002; Lebiere, 1999) and a spiking neural model of simple mental addition—which was later expanded upon to integrate learning (Aubin et al., 2017)—was included in Spaun (Choo, 2018; Eliasmith et al., 2012), we are not aware of any models of mental multiplication that operate in a biologically plausible substrate. Hence, ours is unique in combining psychological plausibility with an anatomical mapping and spiking neural implementation.

We proceed as follows: first, we summarize the human mental multiplication literature, informing our approach; second, we describe our model's operation, based on repeated addition and rule use, and implementation; and, third, we present results of experiments on our model, highlighting replicated human performance levels, including that of both children and adults, and trends, such as problem-size and outlier effects.

Background

Solution Strategies

Retrieval The primary strategy used by children and adults alike to perform mental multiplication is retrieval (i.e., fact recall from memory; Campbell & Graham, 1985; Cooney et al., 1988; Jerman, 1970; Koshmider & Ashcraft, 1991; LeFevre et al., 1996; Lemaire & Siegler, 1995; Siegler, 1988). Only when retrieval fails are other *backup* strategies invoked (Cooney et al., 1988). Retrieval is likely realized in the brain as a learned hetero-associative memory network that maps each multiplication problem to its corresponding solution (Campbell & Graham, 1985). As with other learned behaviours, retrieval is used more often with practice and, thus, age (Cooney et al., 1988; Lemaire & Siegler, 1995; Siegler, 1988). However, before retrieval can succeed, other strategies must be used to facilitate memorization.

Repeated Addition Mathematically, multiplication is defined as the addition of a multiplicand the number of times of a multiplier (e.g., $4 \times 3 = 4 + 4 + 4 = 12$, where 4 is the multiplicand and 3 is the multiplier). This is typically how the operation is taught to children (Jerman, 1970), making mentally multiplying with repeated addition common (Cooney et al., 1988; LeFevre et al., 1996; Lemaire & Siegler, 1995; Siegler, 1988). This strategy reliably produces correct answers when applied carefully, but its use is generally avoided because it requires significant mental effort and time (LeFevre et al., 1996; Lemaire & Siegler, 1995); humans prefer to simply retrieve memorized answers (Cooney et al., 1988). However, even as humans learn shortcuts to bypass repeated addition, this cumbersome procedure remains the main backup strategy invoked when retrieval fails (Lemaire & Siegler, 1995; Siegler, 1988). Repeated addition is thus fundamental to mental multiplication skill (Cooney et al., 1988), both preceding and acting as a fail-safe for memorization. Some have argued that, when executing repeated addition, humans tend to treat the smaller factor as the multiplicand and the larger factor as the multiplier (Cooney et al., 1988; Jerman, 1970), but others have concluded the opposite (Lemaire & Siegler, 1995). Given this disagreement, we take neither view as universally correct. Instead, following Siegler (1988), our model treats the first factor as the multiplicand and the second factor as the multiplier; note that, because we test both permutations of each factor combination, this equates to random role assignment.

Rule Use Simple tricks can be used to solve special types of multiplication problems (Cooney et al., 1988; LeFevre et al., 1996), such as multiplication-by-0 (because $n \times 0 = 0$) and multiplication-by-1 (because $n \times 1 = n$). Distinguishing rule use from retrieval in psychological experiments may be difficult, but the underlying mechanisms are distinct; rule use is scalable and generalizes to arbitrary inputs, while retrieval requires prior experience with the presented problem. That humans are able to correctly and rapidly solve never-before-seen problems conforming to the structure of a rule (e.g., $9,729 \times 0$ or $8,771 \times 1$) suggests underlying rule use.

Summary Retrieval, repeated addition, and rule use are the strategies behind the vast majority of mental multiplication in humans. Directly comparing the effectiveness of different strategies is difficult due to selection bias (Siegler & Lemaire, 1997), but children (Jerman, 1970; Lemaire & Siegler, 1995; Siegler, 1988) and adults (LeFevre et al., 1996) alike choose different strategies for different problems. With practice, humans both improve at using their favourite strategies (Jerman, 1970) and evolve the means by which they select strategies (Lemaire & Siegler, 1995). The general trend observed is a shift from *manual* methods (e.g., repeated addition) to *automatic* methods (e.g., retrieval) over time (Cooney et al., 1988; Lemaire & Siegler, 1995; Siegler, 1988). To summarize, we model humans as using following method to perform mental multiplication: If they have seen the presented problem before and are confident in their memorized solution, they retrieve this solution; otherwise, they manually compute the solution using a backup strategy. If a rule can be applied, they apply one; otherwise, they perform repeated addition.

Performance Patterns

Problem-Size Effect As with the analogous addition operation (Zbrodoff & Logan, 2005), the accuracy of human solutions to multiplication problems decreases and response time increases as the value of the correct product increases. Both children (Campbell & Graham, 1985; Siegler, 1988) and adults (Campbell & Graham, 1985; Koshmider & Ashcraft, 1991; LeFevre et al., 1996) exhibit this effect. This trend points to the use of repeated addition; incrementing to larger products should take longer and introduce more errors than incrementing to smaller products.

Outliers Certain types of multiplication problems are consistently solved more accurately and faster than others. Primary examples are the aforementioned multiplication-by-0 and multiplication-by-1 types (Cooney et al., 1988; Jerman, 1970; LeFevre et al., 1996); this suggests humans use an alternative strategy, rule use, to solve these problems. Other examples, such as multiplication-by-5 and ties (Jerman, 1970; Siegler, 1988), may be attributable to retrieval processes; these problems tend to be solved, and therefore reinforced in memory, more often than other arbitrary products, leading to subsequent more accurate and faster recall. Our model addresses the former two outliers, but not the latter two.

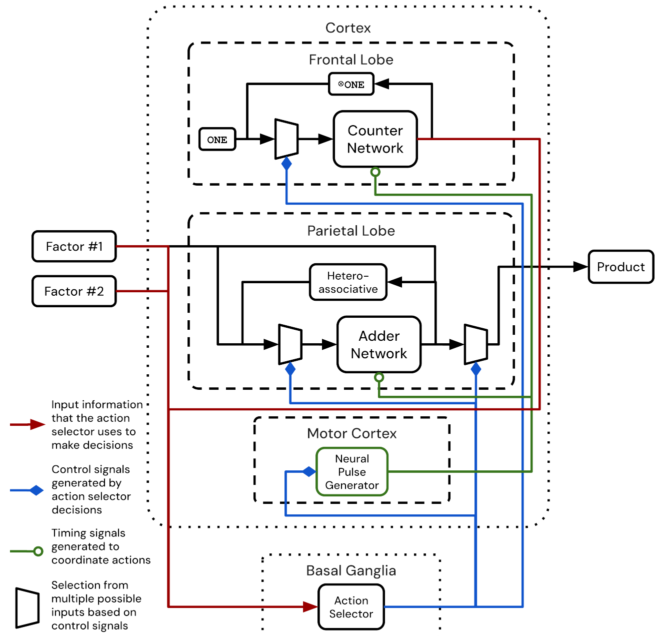


Figure 1: Model architecture, as mapped onto anatomical correlates (Choo, 2018; Ischebeck et al., 2006).

Model

The scope of our model (Fig. 1) is to implement two of the key strategies humans use to perform mental multiplication: repeated addition and rule use. In other words, our model captures the computations carried out by the backup networks that the brain resorts to using when an answer cannot be produced via retrieval. In this sense, we model how a learning child or forgetful adult might perform mental multiplication.

Although we believe our method can scale to multiply ever larger numbers, we constrain all input factors to be < 10 in this work. This domain is a natural choice because it matches the conventional times table taught to children and, thus, is the focus of the experimental literature we compare to. Additionally, solving larger problems with the same strategies would be psychologically implausible. This is because multiplying multi-digit factors is substantially more difficult, reaching the limits of mental arithmetic and requiring complex strategies that often rely on tools outside of the brain (Fuson, 2020).

Repeated Addition Previously, mental addition has been modelled as an iterative counting procedure (Choo, 2018;

Eliasmith et al., 2012) leading to eventual memorization (Aubin et al., 2017). This is because counting—incrementing a natural number by 1—is the fundamental structured numerical operation that human brains seem capable of performing (Groen & Parkman, 1972). To construct a system capable of adding two arbitrary natural numbers using only a simple increment-by-1 operation, extra machinery is needed. Aubin et al. (2017) used two working memories—one holding a periodically-recalculated cumulative sum (the *accumulator*) and the other holding the number of counting operations completed thus far (the *counter*)—that increment in parallel. The accumulator starts from the value of one of the addends, while the counter starts from 0. Once the counter reaches the value of the other addend, the accumulator holds the sum of the two addends, so the computation halts. This construction enables their model to compute addition manually. However, as with multiplication, humans shift from using this manual counting-based strategy to using an automatic retrieval-based strategy over time (Siegler, 1987). Our model assumes the agent performing the mental multiplication has already mastered addition, meaning that the agent has memorized all the basic sums and can therefore perform addition by simply querying a hetero-associative memory. In this sense, we assume a network like the one developed by Aubin et al. (2017) already exists. Children typically learn multiplication several years after they have learned addition (Stokke, 2015), so this assumption is plausible.

Our model (Fig. 1) employs an iterative mechanism similar to the one proposed for addition. Before, addition was realized as iterative counting; now, multiplication can be realized as iterative addition. To construct a similar system capable of multiplying two arbitrary natural numbers using the repeated addition algorithm, we replace the increment operation applied at each iteration of the counting-based addition algorithm with an addition operation computed as retrieval of memorized facts. Like for addition, this system works by having two working memories—one holding a periodically-recalculated cumulative product (the *adder*) and the other holding the number of addition operations completed thus far (the *counter*)—operate in parallel. The adder starts from the value of the multiplicand, while the counter starts from 1. The multiplicand is periodically added to the cumulative product stored in the adder. This addition operation is performed by querying the aforementioned hetero-associative memory with the value of the current cumulative product and the value of the multiplicand; the retrieved response to this query then becomes the new cumulative product held in the adder. The counter is periodically incremented in parallel with the operation of the adder network. Once the counter reaches the value of the multiplier, the adder holds the product of the two factors, so the computation halts. This construction enables our model to compute multiplication via repeated addition.

Rule Use Repeated addition works for arbitrary factors, but is unnecessary for problems with a 0 or 1 factor. These problems are best solved by applying rules (i.e., $n \times 0 = 0$ and $n \times 1 = n$) that simply map special problem types to standard products without the need for a time-consuming calculation. If either factor is 0, then our model sends 0 to output; if either factor is 1, then our model routes the value of the other factor to output. Only when neither of these conditions are met does our model execute the repeated addition strategy.

Implementation

We built our model with the Neural Engineering Framework (NEF; Eliasmith & Anderson, 2003) and implemented it in spiking neurons with the Nengo neural simulator (Bekolay et al., 2014). Our model uses the Semantic Pointer Architecture (SPA; Eliasmith, 2013), which implements a Vector Symbolic Algebra (VSA), to generate a structured cognitive representation of numbers, and uses a basal ganglia model to support cognitive rule following (Stewart et al., 2010).

Our model defines numbers as semantic pointers (SPs), symbols representable as high-dimensional vectors. These SPs are constructed recursively so that each number's SP is simply the preceding number's SP bound with a fundamental ONE SP: $\texttt{NUMBER}_{N+1} = \texttt{NUMBER}_N \otimes \texttt{ONE}$ (Choo, 2010). Our work demonstrates that these techniques can support a psychologically plausible model of mental multiplication.

Results

Experimental Setup

To compare our results with available human data, we built two versions of our model: a less accurate, slower *child* model and a more accurate, faster *adult* model. The child model was tuned to count slowly and generated with few resources, while the adult model was tuned to count quickly and generated with substantial neural resources. These resources model the increased number of neurons and synapses dedicated to the mental multiplication task as skill develops. We explored the performance of each by adjusting the number of neurons and SP vector dimensions present in the model. We tested neuron counts per dimension, n_n, of 100, 150, and 200 for the child model and 200, 300, and 400 for the adult model, as well as vector dimensionalities, n_d, of 32 and 64 for the child model and 128 and 256 for the adult model. Over these ranges, the smallest child model comprised $231,790$ neurons and the largest adult model comprised $1,936,890$ neurons. All 12 configurations were tested for 15 trials, each with a unique random initialization, requiring 180 experiments.

We tested our model on each possible multiplication problem of interest (i.e., all combinations of factors ranging from 0 through 9), forming a testing dataset of 100 samples. Because different mechanisms were used to solve different problems, we segregated performance by solution strategy; the testing dataset was divided into problems solved by repeated addition (64 samples of the form $n \times m$ where $2 \leq n, m \leq 9$) and problems solved by rule use (36 samples of the form $n \times 0$, $0 \times n$, $n \times 1$, or $1 \times n$).

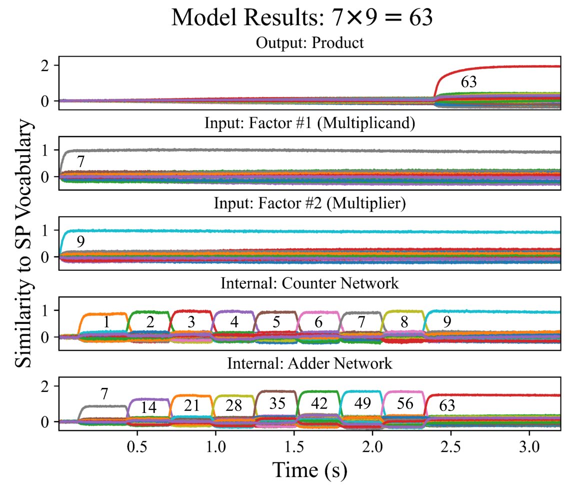
We analyzed two metrics: *accuracy*, the correctness of the computed product, and *latency*, the time taken to compute it. To measure accuracy, we compared the computed product (i.e., the SP most similar to the value represented by the neural activity in our model's output working memory) to the correct product. To measure latency, we extracted the simulation time elapsed before the problem was solved (i.e., the point in simulation time when this SP was first produced).

Model Performance

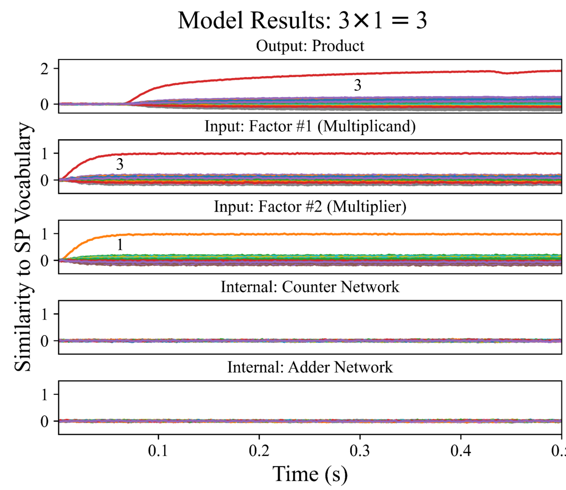
Qualitatively, our model behaved well (Fig. 2). Both the child and adult models were able to properly execute both the repeated addition and rule use strategies. In an example of repeated addition (Fig. 2a), the counter cycled through representations of the `ONE`, `TWO`, `THREE`, ..., and `NINE` SPs while the adder cycled through representations of the `SEVEN`, `FOURTEEN`, `TWENTY_ONE`, ..., and `SIXTY_THREE` SPs using the retrieval-based addition network; once the counter's represented SP matched the SP represented in the second factor's working memory, `NINE`, the basal ganglia routed the adder's represented SP, `SIXTY_THREE`, to the output working memory. In an example of rule use (Fig. 2b), the basal ganglia detected that the SP represented in the second factor's working memory, `ONE`, rendered the problem rule-eligible and routed the SP represented in the first factor's working memory, `THREE`, to the output working memory accordingly; no counting operations were initiated. In an example of model failure (Fig. 2c), the representational quality in the adder degraded over time due to a lack of resources (i.e., too few neurons), causing no SP to be produced in the output working memory.

Quantitatively, our model also behaved well (Table 1). Model performance scaled directly with both neuron count and dimensionality (Fig. 3), meaning our model could be tuned to match the accuracy of humans of various skill levels. To understand how well our model captured human performance and behaviour, we directly compared our results to results from four studies of humans (Cooney et al., 1988; LeFevre et al., 1996; Lemaire & Siegler, 1995; Siegler, 1988) that break down detailed data by solution strategy (Table 2).

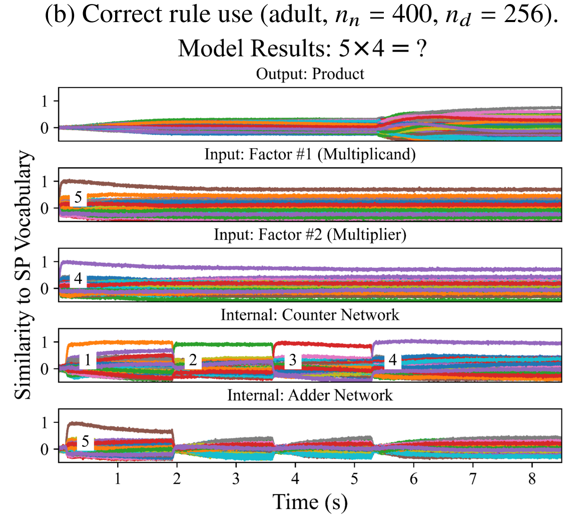
Child Model Overall, data from the child model closely matched data from studies of children. For repeated addition, the child model's average accuracy of 57.9% fell within the experimental range of 52.3%, 53%, and 59%, but its average latency of 6.795s was lower than the experimental range of 11.8s and 23.3s. Likewise, for rule use, the child model's average accuracy of 99.9% was similar to the experimental 98.3%, but its average latency of 0.322s had no empirical reference point. Model accuracy varied between trials due to effects of random initialization, but degraded gracefully as model resources deteriorated. Model latency appeared to decrease with reduced resources, but this trend arose merely because latency figures were aggregated only among problems solved correctly, consistent with much of the experimental literature, and problems that required less time to solve were more likely to be solved correctly. Importantly, the two main performance patterns observed in children—problem-size and outlier effects—were qualitatively replicated.



(a) Correct repeated addition (adult, $n_n = 200$, $n_d = 128$).



(b) Correct rule use (adult, $n_n = 400$, $n_d = 256$).



(c) Incorrect repeated addition (child, $n_n = 50$, $n_d = 64$).

Figure 2: Selected sample model results under various configurations. The rows of each figure show the similarity of the neural activity in each of the model's working memories to all number SPs throughout the course of the simulation, with the labels indicating the SP with the maximum similarity.

Table 1: Model performance, in terms of both accuracy and latency, on the testing dataset, and each constituent portion, under each configuration (see also Fig. 3). Metrics are the mean value and standard deviation across 15 trials. Latency figures include the standard 85ms for visual processing (Anderson & Lebiere, 2014) and 150ms for motor processing (Meyer & Kieras, 1997).

n_n	n_d	Total (n = 100)		Repeated Addition (n = 64)		Rule Use (n = 36)	
		Accuracy (%)	Latency (s)	Accuracy (%)	Latency (s)	Accuracy (%)	Latency (s)
100	32	64.1 ± 19.8	2.762 ± 1.562	44.0 ± 30.9	6.468 ± 2.691	99.8 ± 0.7	0.321 ± 0.002
100	64	68.3 ± 19.3	3.233 ± 1.755	50.4 ± 30.2	6.460 ± 2.657	100.0 ± 0.0	0.322 ± 0.001
150	32	73.6 ± 21.4	3.525 ± 2.099	58.9 ± 33.5	6.335 ± 3.027	99.8 ± 0.7	0.321 ± 0.002
150	64	82.0 ± 16.6	4.157 ± 1.664	71.9 ± 25.9	7.551 ± 2.283	100.0 ± 0.0	0.322 ± 0.002
200	32	65.2 ± 24.6	3.114 ± 2.228	45.7 ± 38.4	5.644 ± 3.807	99.8 ± 0.7	0.321 ± 0.002
200	64	84.9 ± 17.3	4.789 ± 1.362	76.5 ± 27.0	8.309 ± 0.818	100.0 ± 0.0	0.322 ± 0.001
Child Model Average		73.0	3.597	57.9	6.795	99.9	0.322
200	128	96.0 ± 11.1	1.198 ± 0.068	93.8 ± 17.3	1.749 ± 0.093	100.0 ± 0.0	0.323 ± 0.002
200	256	96.6 ± 12.1	1.188 ± 0.125	94.7 ± 18.9	1.719 ± 0.038	100.0 ± 0.0	0.325 ± 0.001
300	128	96.7 ± 7.7	1.219 ± 0.039	94.9 ± 12.1	1.759 ± 0.040	100.0 ± 0.0	0.323 ± 0.001
300	256	96.8 ± 12.1	1.204 ± 0.122	95.0 ± 18.9	1.744 ± 0.020	100.0 ± 0.0	0.325 ± 0.002
400	128	96.9 ± 8.5	1.223 ± 0.041	95.2 ± 13.3	1.764 ± 0.046	99.8 ± 0.7	0.323 ± 0.001
400	256	96.5 ± 11.9	1.205 ± 0.119	94.6 ± 18.5	1.747 ± 0.022	100.0 ± 0.0	0.325 ± 0.001
Adult Model Average		96.6	1.206	94.7	1.747	100.0	0.324

To analyze the problem-size effect, we determined the relationship between child model performance, in terms of both accuracy and latency, and problem size, in terms of both the value of each factor and the correct product, for the repeated addition strategy (Fig. 4). To facilitate comparison, we aggregated data from all model instantiations, just as analyses in studies of humans aggregate data across all participants. No correlation existed between model accuracy (ρ = 0.056) or latency (ρ = 0.001) and the value of the first factor, but significant correlations existed between both model accuracy (ρ = −0.924) and latency (ρ = 0.999) and the value of the second factor. This is because the value of the second factor exclusively determined the number of counting and addition operations completed in the algorithm, and each additional operation introduced another systematic delay. This strong linear relationship between model latency and the value of the second factor matched the linear delay pattern observed when humans count (Cordes et al., 2001), indicating psychological plausibility. Regression analysis on the measured latencies revealed an average operation delay of 1.687s. Each additional operation also introduced possibility for error, producing the significant manifestation of the problem-size effect in model accuracy. Additionally, as a byproduct of the second factor, significant correlations existed between both model accuracy (ρ = −0.571) and latency (ρ = 0.679) and value of the correct product. Siegler (1988) observed a similar accuracy-product correlation in children (ρ = −0.836) between our accuracy-multiplier (ρ = −0.924) and accuracy-product (ρ = −0.571) correlations; however, by regression coefficient, the effect was stronger in children (β = −0.966) than in our model (β = −0.309). This qualitative replication of the problem-size effect supports psychological plausibility.

Table 2: Model performance, in terms of both accuracy and latency, compared with relevant human data. Note the following: for Cooney et al. (1988), we aggregate results from both grade levels and include non-responses; for Lemaire and Siegler (1995), we report results from the end of the school year and again include non-responses; and, for LeFevre et al. (1996), we approximately aggregate results from both experiments and include the number series strategy with repeated addition.

Study	Subjects	Repeated Addition		Rule Use	
		Accuracy (%)	Latency (s)	Accuracy (%)	Latency (s)
Cooney et al. (1988)	American Third/Fourth-Graders	52.3	—	98.3	—
Lemaire and Siegler (1995)	French Second-Graders	53	11.8	—	—
Siegler (1988)	American Third-Graders	59	23.3	—	—
Ours	**Child Model Average**	**57.9**	**6.795**	**99.9**	**0.322**
LeFevre et al. (1996)	Canadian Undergraduates	97	1.493	99	0.881
Ours	**Adult Model Average**	**94.7**	**1.747**	**100.0**	**0.324**

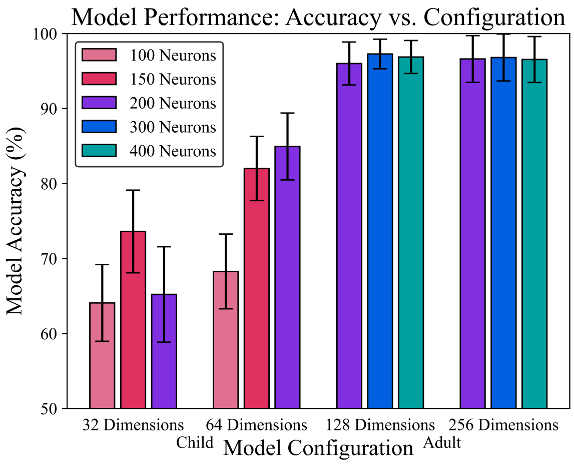


Figure 3: Model performance, in terms of accuracy, under each configuration (see also Table 1). Bars are the mean across all trials and ranges are the standard error of the mean.

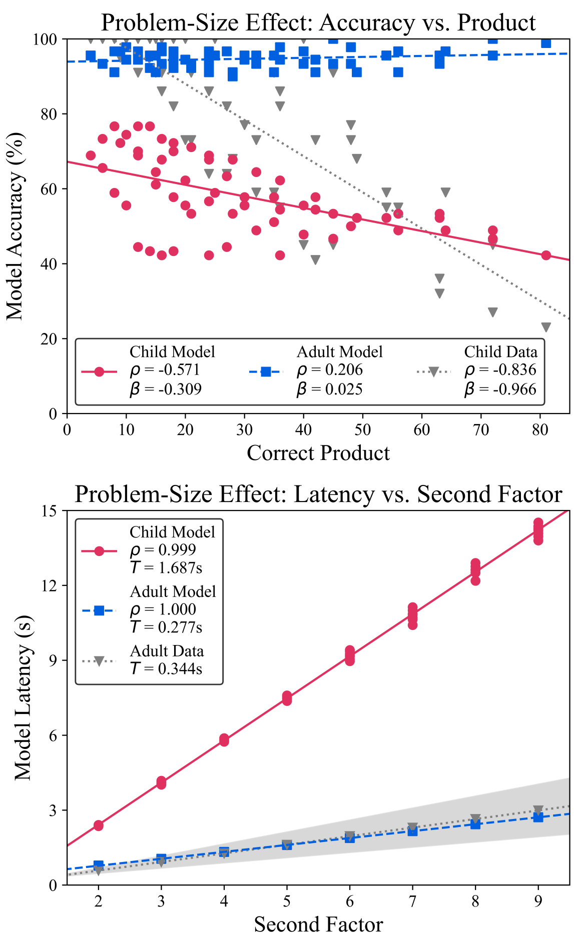


Figure 4: Model problem-size effect, in terms of both accuracy and latency, when using repeated addition. Recall the experimental accuracy-product correlation ($\rho = -0.836$) and coefficient ($\beta = -0.966$) for children (Siegler, 1988). Recall also the linearity of human counting (Cordes et al., 2001) and adult counting delay of 344±135ms (Landauer, 1962).

To analyze outliers, we contrasted child model performance on the different problem types. Multiplication-by-0 and multiplication-by-1 problems were solved much more accurately (1.7×) and much faster (21.0×) than others because our model applied the rule use strategy instead of the repeated addition strategy. This disparity was particularly evident as model resources deteriorated; even as repeated addition's accuracy dropped to 0.0% in some trials with extremely limited resources, rule use's accuracy remained perfect in the vast majority of trials. This qualitative replication of outlier effects also supports psychological plausibility.

Adult Model Overall, data from the adult model closely matched data from studies of adults. For repeated addition, the adult model's average accuracy of 94.7% and latency of 1.747s were similar to the experimental accuracy of 97% and latency of 1.493s. Likewise, for rule use, the adult model's average accuracy of 100.0% and latency of 0.324s were similar to the experimental accuracy of 99% and latency of 0.881s. Other than one outlier trial, models performed consistently well. The adult model scored 100.0% accuracy in 76.7% of trials, indicating that our model is capable of perfect performance.

Analyzing problem-size and outlier effects in the adult model was difficult because few errors were made. However, some important patterns endured, again supporting psychological plausibility. As in the child model, there was a strong linear relationship between latency and the value of the second factor ($\rho = 1.000$) that matched observations of humans (Cordes et al., 2001). Additionally, the adult model had an average operation delay of 277ms that matched the human counting delay of 344±135ms (Landauer, 1962). Moreover, multiplication-by-0 and multiplication-by-1 problems were solved more accurately (1.1×) and faster (5.4×) than others.

Conclusion

In this paper, we presented a novel model of mental multiplication. Our model realized a psychologically plausible hypothesis of how humans, particularly inexperienced children and forgetful adults, might mentally multiply small natural numbers by implementing repeated addition and rule use, two strategies commonly used by humans. Our model was implemented in biologically plausible spiking neurons with naturally tunable performance, allowing it to achieve perfect accuracy. This construction enabled our model to match various human performance levels, including that of both children and adults, and qualitatively replicate human performance patterns, such as problem-size and outlier effects. This resemblance to humans suggests that our methods may accurately capture features of cognition, shedding light on the potential processes behind this task by further evidencing repeated addition and rule use in humans. To the best of our knowledge, ours is the only spiking neural model of mental multiplication; as such, it provides a tangible example of how structured, rule-based reasoning may be realized in biological systems. In the future, our model may be extended to integrate the use of retrieval with an accompanying learning process.

Acknowledgements

The authors would like to thank Xuan Choo for discussions that helped improve this paper. This work was supported by CFI (52479-10006) and OIT (35768) infrastructure funding as well as the Canada Research Chairs program, NSERC Discovery grant 261453, AFOSR grant FA9550-17-1-0644 and NSERC CGS-M graduate funding.

References

Anderson, J., & Lebiere, C. (2014). *The atomic components of thought*. Taylor & Francis.

Aubin, S., Voelker, A. R., & Eliasmith, C. (2017). Improving with practice: A neural model of mathematical development. *Topics in Cognitive Science*, *9*(1), 6–20.

Bekolay, T., Bergstra, J., Hunsberger, E., DeWolf, T., Stewart, T., Rasmussen, D., Choo, X., Voelker, A., & Eliasmith, C. (2014). Nengo: A Python tool for building large-scale functional brain models. *Frontiers in Neuroinformatics*, *7*(48), 1–13.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (Pan troglodytes). *Journal of Comparative Psychology*, *103*(1), 23–31.

Brannon, E. M. (2005). What animals know about numbers. In *The handbook of mathematical cognition* (pp. 85–107). Psychology Press.

Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology*, *39*(2), 338–366.

Cantlon, J. F., & Brannon, E. M. (2007). Basic math in monkeys and college students. *PLoS Biology*, *5*(12).

Choo, X. (2010). *The ordinal serial encoding model: Serial memory in spiking neurons* [Master's Thesis]. University of Waterloo.

Choo, X. (2018). *Spaun 2.0: Extending the world's largest functional brain model* [PhD thesis]. University of Waterloo.

Cooney, J. B., Swanson, H. L., & Ladd, S. F. (1988). Acquisition of mental multiplication skill: Evidence for the transition between counting and retrieval strategies. *Cognition and Instruction*, *5*(4), 323–345.

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

Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. Oxford University Press.

Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. MIT Press.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., & Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science*, *338*(6111), 1202–1205.

Fuson, K. (2020). The best multidigit computation methods: A cross-cultural cognitive, empirical, and mathematical analysis. *Universal Journal of Educational Research*, *8*, 1299–1314.

Griffiths, T. L., & Kalish, M. L. (2002). A multidimensional scaling approach to mental multiplication. *Memory & Cognition*, *30*(1), 97–106.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, *79*(4), 329–343.

Hendrycks, D., Burns, C., Kadavath, S., Arora, A., Basart, S., Tang, E., Song, D., & Steinhardt, J. (2021). Measuring mathematical problem solving with the MATH dataset. *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*.

Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstätter, F., Benke, T., Felber, S., & Delazer, M. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. *NeuroImage*, *30*(4), 1365–1375.

Jerman, M. (1970). Some strategies for solving simple multiplication combinations. *Journal for Research in Mathematics Education*, *1*(2), 95–128.

Koshmider, J. W., & Ashcraft, M. H. (1991). The development of children's mental multiplication skills. *Journal of Experimental Child Psychology*, *51*(1), 53–89.

Landauer, T. K. (1962). Rate of implicit speech. *Perceptual and Motor Skills*, *15*(3).

Lebiere, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft*, *8*(1), 5–19.

LeFevre, J.-A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General*, *125*(3), 284–306.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, *124*(1), 83–97.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part i. basic mechanisms. *Psychological Review*, *104*(1), 3–65.

Ouellette, S., Pfister, R., & Jud, H. (2024). Counting and algorithmic generalization with transformers. https://arxiv.org/abs/2310.08661

Schön, M., Ebner, M., & Kothmeier, G. (2012). It's just about learning the multiplication table. *ACM International Conference Proceeding Series*, 73–81.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*(3), 250–264.

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General*, *117*(3), 258–275.

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of

ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, *126*(1), 71–92.

Steel, S., & Funnell, E. (2001). Learning multiplication facts: A study of children taught by discovery methods in england. *Journal of Experimental Child Psychology*, *79*(1), 37–55.

Stewart, T. C., Choo, X., & Eliasmith, C. (2010). Dynamic behaviour of a spiking model of action selection in the basal ganglia. *10th International Conference on Cognitive Modeling*.

Stokke, A. (2015). What to do about Canada's declining math scores. *C.D. Howe Institute Commentary*, (427), 1–15.

Yehudai, G., Kaplan, H., Ghandeharioun, A., Geva, M., & Globerson, A. (2024). When can transformers count to n? https://arxiv.org/abs/2407.15160

Yuan, Z., Yuan, H., Tan, C., Wang, W., & Huang, S. (2023). How well do large language models perform in arithmetic tasks? https://arxiv.org/abs/2304.02015

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem-size effect. In *The handbook of mathematical cognition* (pp. 331–345). Psychology Press.