Evaluating the Evidence Base for Rocket Math: A Science of Learning Translational Analysis Using the Donoghue and Horvath Framework

This report was submitted to The University of Melbourne as part of the Fundamentals of the Science of Learning subject in June 2026

by Ash Abdou

Abstract

Rocket Math is a widely used mathematics intervention designed to develop automaticity in basic number facts through structured, daily retrieval practice. This paper evaluates the evidence supporting Rocket Math using the Donoghue and Horvath translational framework, which classifies research according to its proximity to classroom practice. Evidence is synthesised to examine three key claims of the program: that rapid responses indicate automaticity, that daily retrieval practice promotes long-term retention, and that structured peer practice enhances learning. The analysis demonstrates that while no single study directly validates Rocket Math as a complete intervention, converging evidence across the translational continuum provides strong support for its core instructional principles. This synthesis highlights how findings from cognitive neuroscience, experimental psychology, and classroom research collectively justify the program's emphasis on retrieval practice, frequent rehearsal, and fluency development. The paper also illustrates the value of translational science as a framework for evaluating educational interventions and informing evidence-based teaching practice.

Across the Australian Curriculum, fluency is recognised as an important component of mathematical proficiency (Australian Curriculum, Assessment and Reporting Authority, 2024). Research suggests that fluency begins with manipulating numbers in the early years, forming a critical foundation for later mathematical competence (Geary, 2011). As students develop this foundation, they improve their ability to recall basic number facts. This in turn can free working memory resources, improving learning of more complex problems (Cartwright, 2022; Dole et al., 2018; Miller et al., 2011; Yu et al., 2025).

One structured approach to improving recall with basic addition, subtraction, multiplication and division is the Rocket Math Worksheet Program (Crawford, n.d.). The program aims to develop automatic recall through structured daily peer-practice, during which students take turns recalling facts aloud for two to three minutes, while their partner provides immediate corrective feedback. Progress through the sequence is determined by a one-minute timed assessment where students advance by meeting a predetermined goal.

This paper evaluates Rocket Math within an Australian primary school setting, focusing on students aged 6-10. The aim of Rocket Math is for students to recall basic number facts automatically, without relying on procedural strategies such as counting. According to its creator, Don Crawford, this is called automaticity, and “automaticity is the point of Rocket Math” (Crawford, 2015a). This report evaluates two claims that underpin this theoretical rationale:

1. Recalling number facts within 2–3 seconds is an indicator of automaticity.
2. Automaticity is achieved through structured daily peer-practice.

The extent to which these claims are supported by evidence across neuroscience, cognitive psychology, and classroom-based research, forms the focus of this report.

The Layered abstraction framework for the learning sciences

The layered abstraction framework, by Donoghue and Horvath (2016) categorises research associated with different disciplines into five layers that are differentiated by varying degrees of biological complexity. It will be used to enable research to be interpreted within the discipline most closely associated with each layer and help avoid incompatible applications of findings (Donoghue & Horvath, 2016). As the Rocket Math claims are most closely associated with neurological, psychological and sociological disciplines, this report focuses exclusively on research from Layers III, IV and V of the framework.

Table 1. Layered abstraction framework

Layer	Layer name	Relevant discipline	Function
V	Sociocultural	Education, social sciences	Learning occurs through social interaction which create shared community expectations and in turn drive decisions about education, such as what is taught and how it is taught.
IV	Individual	Cognitive & behavioural psychology	When people interact, we can observe patterns in behaviour. Learning can be observed and measured through changes in individual performance, encompassing biological, psychological and emotional systems.
III	Cerebral	Systems, cognitive and behavioural neurosciences	The brain works through networks of connected cells and the activity of these networks allows information to be processed often automatically and without awareness.
II	Cellular	Biology/pure neuroscience	Individual cells, particularly neurons, can store, encode, process and send information through chemical signals. Changes inside these cells form the biological foundation of learning.
I	Physical	Physics	Information can be stored and processed in physical systems like complex molecules or atoms. This forms the most fundamental level of learning in the framework, but it has little direct relevance to classroom teaching.

Note. adapted from Donoghue & Horvath (2016)

The translation schema

The translation schema governs how findings can be interpreted and applied across layers. It is particularly important when synthesising evidence and drawing conclusions as it provides principles for determining whether findings from one layer can be meaningfully translated to another (Donoghue & Horvath, 2016). The key features of the translation schema most relevant to this report are summarised below.

Downward compatibility

Findings established within a higher layer must remain consistent with what is known about learning at lower layers (Donoghue & Horvath, 2016). For example, classroom strategies at Layer V should align with evidence from the Individual layer and subsequent lower layers.

Upward unpredictability

Findings established at lower layers may not necessarily be extrapolated to higher layers (Donoghue & Horvath, 2016). For example, evidence from Cellular level research cannot alone determine the effectiveness of classroom instructional strategies at Layer V. Evidence in lower layers, may be useful, however, in generating a hypothesis about learning in higher layers.

Translational contiguity

The framework requires that research findings be translated sequentially across adjacent layers of analysis (Donoghue & Horvath, 2016). For research in the lower layers to be meaningful at Layer V, it must be translated through adjacent layers of analysis. Skipping layers is likely to result in false conclusions about the impact and relevance of phenomena in lower layers.

Analysis of claims

Claim 1

To evaluate claim 1, it is necessary to examine what evidence is used to distinguish automatic recall from the use of procedural strategies such as counting.

A study by Stickney et al. (2012), used the computerised assessment system MathFacts in a Flash (Renaissance Learning, 2003) to assess 89,159 students from Grades 1–3 across the United States (US). Researchers investigated how many students demonstrated automaticity in basic addition and subtraction involving numbers between 0-10, and multiplication and division involving numbers between 0-12. Using timed multiple-choice tasks, the assessment measured both “accuracy and speed of student retrieval of math facts” (Stickney et al., 2012, p.85).

Stickney et al. define automaticity as “the ability to deliver a correct answer immediately from memory without conscious thought, as opposed to relying on calculation”, and further suggest that it is demonstrated by “a recall rate of 2 to 3 seconds”. This interpretation aligns closely with Crawford’s assertion that “3 seconds to input an answer means students don’t have time to count on their fingers or to ‘figure out’ a fact – they just have to remember it” (Crawford, 2020a).

The study's relevance to Rocket Math is somewhat limited because retrieval was measured using a computerised task rather than a worksheet task. Consequently, the response times identified may not transfer directly to a classroom context where students write and verbalise responses. The findings should be interpreted cautiously, as the study did not include a control group or random assignment, meaning that the researchers were “limited in drawing conclusions about causation” (Stickney et al., 2012, p. 92). Nevertheless, the consistency of the findings across a sample of 89,159 students lends credibility to the reported relationship between recalling number facts within 2–3 seconds and automaticity.

The research was conducted at a classroom level, however, because the primary focus of the study was the measurement of individual student’s retrieval speed, the study is best situated within Layer IV of the Donoghue and Horvath framework.

While Stickney et al. support the claim that recall within 2–3 seconds is an indicator of automaticity, research by Díaz-Barriga Yáñez et al. (2023) challenges the assumption that rapid response times reflect the ability to deliver an answer without procedural calculations.

Their study investigated whether improvements in children's performance on single-digit addition problems (sums ≤ 10) reflected a shift from counting strategies to direct memory retrieval, or whether counting itself became more efficient. Participants spanned four age groups: 8-9, 11-12, 14-15 and adults with a mean age of around 23. They completed a vocal task, where they said answers aloud and then a silent task inside a functional magnetic resonance imaging (fMRI) scanner, where they mentally solved problems and pressed a button to answer. The study found that brain regions associated with processing increasingly difficult addition problems remained active from childhood through to adulthood. Key regions included the intraparietal sulcus, medial prefrontal cortex and insula. The authors point out that these regions are not typically associated with automatic memory retrieval, which would predict activity in left temporo-parietal cortex instead (Vogel and De Smedt, 2021 as cited in Díaz-Barriga Yáñez et. al., 2023). The continued activation of procedural regions across development is therefore difficult to reconcile with the view that fast performance simply reflects direct memory retrieval (Díaz-Barriga Yáñez et. al., 2023).

Of the four age groups, the 8–9-year-olds are the most similar those using Rocket Math. Researchers found that these children took progressively longer to solve addition problems as the size of the problem increased. This phenomenon, known as the problem-size effect, refers to the finding that people take longer to solve addition problems as the sum of the operands increases (Díaz-Barriga Yáñez et. al., 2023). For example, 1+2 is solved faster than 2+3. For the 8–9-year-old group, researchers measured how much response times increased for each unit increase in problem sum - for example, from a sum of 3 to 4. Researchers found that response times increased by approximately 157 milliseconds per unit increase in the vocal task and 202 milliseconds in the fMRI task. For older participants, this rate of increase became smaller, suggesting that the procedures used to solve these problems became more efficient with development and practice (Díaz-Barriga Yáñez et. al., 2023).

The study's relevance to Rocket Math's broader claims is limited by its focus on addition alone, leaving open the question of whether similar patterns would be observed for subtraction, multiplication and division. Furthermore, as all participants were native

French speakers, the findings may not fully generalise across different linguistic and educational contexts.

Rocket Math's claim that recall within 2–3 seconds demonstrates automaticity rests on the assumption that responses are produced either through direct memory retrieval or via procedural strategies. Díaz-Barriga Yáñez et al. provide Layer III evidence that fast, automatic performance may not reflect pure memory retrieval, but rather an exceptionally fast procedural process that has become so efficient, it is no longer consciously accessible. The authors themselves acknowledge this complexity, noting "it is difficult to distinguish automatized counting from retrieval at the behavioural level" (Díaz-Barriga Yáñez et al., 2023, p. 2). From a classroom perspective, however, the development of efficient and accurate responding is likely to be more important than whether performance is achieved through retrieval or highly automatised procedures.

Considering the principle of translational contiguity, findings at Layer III cannot be directly extrapolated to Layer V without translation through adjacent layers of analysis (Donoghue & Horvath, 2016). Consequently, although Díaz-Barriga Yáñez et al. provide Layer III evidence regarding arithmetic processing, their findings cannot be used to make direct claims about classroom instructional practices. The evidence presented by Stickney et al., however, provides Layer IV evidence, thereby enabling translation toward Layer V and the generation of classroom hypotheses; however, consistent with the principle of upward unpredictability in the Donoghue and Horvath framework, such Layer IV findings cannot directly predict educational outcomes at Layer V. Nevertheless, as automaticity is fundamentally a cognitive process rather than a classroom intervention, Layers III and IV provide the most direct evidence for evaluating claim 1, reducing the need to establish direct support through Layer V research.

Claim 2

Rocket Math suggest that following the recommended practice method (Crawford, 2022) is typically a 'magic bullet' for student progress (Crawford, 2015b). This practice method includes three key components: peer-practice, verbalisation of facts and daily retrieval practice. Each component will be examined to evaluate the extent to which the available evidence supports them.

Peer-practice

A US study, by Kohler et al. (1994) examined the effects of adding peer-practice to teacher-directed mathematics teaching in a Grade 1 classroom. The purpose was to determine whether peer-practice increased participation and improved learning. As a Layer V study, it is relevant to Rocket Math because it examines whether structured peer-practice produces learning gains in a primary classroom setting.

One of the interventions involved students working in pairs for 10-15 minutes to practice skills from topics that included addition and subtraction fact families, coin values, measurement and calendar problems. Although the topics extended beyond number facts, the intervention shares key features with Rocket Math, including paired practice and immediate corrective feedback.

The relevant findings indicate that peer-practice produced substantially larger gains between pre and post testing, with students demonstrating an average improvement of 34%, compared to 15% for teacher-directed instruction. Furthermore, active responding rose to around 69% of observed intervals during peer-practice, compared to approximately 20% during teacher-directed activities, suggesting the peer-practice format substantially increased the number of practice opportunities each student received (Kohler et al., 1994).

The findings regarding increased responding align closely with Rocket Math's emphasis on frequent opportunities to retrieve and rehearse mathematical facts (Kohler et al., 1994). The study did not, however, measure automaticity, and its small sample size (n=18) limits the strength of the conclusions that can be drawn. Furthermore, the peer-practice activities included content such as measurement and calendar problems, meaning the findings may not translate directly to Rocket Math's specific goal of developing automatic recall of number facts. Consequently, caution is required when generalising the results.

Verbalising number facts

Burns et al. (2019) examined interventions designed to improve multiplication fact retention among Year 3 and 4 students with low mathematics achievement. Students practiced multiplication facts through either incremental rehearsal or traditional drill, with facts considered successfully retained if they could be recalled correctly within two seconds the following day. Both interventions required students to verbalise the complete fact and answer aloud. This study provides an opportunity to evaluate Crawford's assertion that "The 'active ingredient,' the thing that makes Rocket Math effective, is verbal rehearsal" (Crawford, 2017).

Because both conditions required students to verbally state facts aloud, differences in outcomes cannot be attributed to oral responding alone. Burns et al. found that incremental rehearsal led to significantly better retention than traditional drill, with students retaining a mean of 3.75 facts compared to 1.79. This suggests that practice structure matters as much as, if not more than, oral responding, and casts doubt on Crawford's claim that verbal rehearsal is the active ingredient in effective fact practice.

This study therefore does not provide conclusive evidence that verbal retrieval is the primary mechanism driving automaticity development, however, given that verbalising facts is just one aspect of the recommended practice method for Rocket Math - the others being peer-practice and daily retrieval - the findings are consistent with oral rehearsal playing a role in multiplication fact retention.

As the research focused on individual cognitive performance, the study operates primarily at Layer IV of the Donoghue and Horvath framework. As such, it provides an important translational bridge between the Layer V findings of Kohler et al., and underlying cognitive processes associated daily peer-practice. This positioning is significant because, according to Donoghue and Horvath's principal of downward compatibility, findings established within a higher layer must remain consistent with what is known about learning at lower layers (Donoghue and Horvath, 2016).

Daily retrieval practice

A study by Breneman-Smith (2024) investigated whether the type and frequency of practice influenced the development of number fact fluency in primary school aged students. The study examined rehearsal and retrieval using daily, second day, and third day schedules. Participants included 196 students across Grades 3-5 from 16 classrooms in seven US schools. Intervention sessions lasted 10 minutes, consisting of two five-minute partner rounds conducted over 45 days. Fluency was measured before and after the intervention using a Mathematics Fluency and Calculation Test (Reynolds, 2015) assessing addition, subtraction, multiplication and division.

Daily practice produced the highest fluency scores, followed by every third day, then every other day (Breneman-Smith, 2024). This finding provides strong support for Rocket Math's recommendation of daily practice. Another relevant finding was that retrieval practice produced significantly stronger gains in mathematical fluency than rehearsal practice. Breneman-Smith, drawing on Brown et al. (2014) and Haebig et al. (2021), point out that retrieval practice strengthens neural pathways associated with long-term retention. In this regard the findings provide support for Crawford's claim that "recalling the answer to a math fact strengthens the neural connection between the problem such as 9 plus 7, and its answer, 16" (Crawford, 2020b).

A limitation of the study was that existing classrooms, rather than randomly assigned students, were allocated to either retrieval practice or rehearsal practice (Breneman-Smith 2024). Because of this, the groups may not have been perfectly equivalent at the start of the study. One class could have been stronger at mathematics, more motivated, or had more prior exposure to the facts being practiced.

Although the study was conducted in classroom settings and examined the impact of practice frequency, which are characteristics of Layer V, it investigated Layer IV variables related to practice type, such as retrieval efficiency. As Donoghue and Horvath (2016) note, "a great deal of psychology is conducted in the Sociocultural Layer V, or at the Layer IV/V interface" (p. 41). Given that the primary outcome of interest was the development of retrieval efficiency, the study is most appropriately classified at Layer IV, albeit situated near the Layer IV/V interface.

Evidence for retrieval practice can also be examined at a lower layer of analysis. Delazer et al. (2003) used fMRI to investigate how repeated retrieval practice alters neural activation patterns during complex arithmetic, situating their research firmly within Layer III of the Donoghue and Horvath framework.

In their study, 13 adult participants were trained over one week on 18 complex two-digit by one-digit multiplication problems and were subsequently scanned while solving both trained and untrained problems. The training was designed to induce a shift from procedural calculation to automatic retrieval, drawing on the claim by Hittmair-Delazer et al. (1994), as cited in Delazer et al. (2003) that “repetition of problems leads to increasing automaticity of retrieval and that the training strategies critically influence the structure of memory representations for arithmetic facts” (p. 77).

Without training, imaging showed activation of a broad fronto-parietal network, particularly the bilateral intraparietal sulcus and inferior frontal gyrus, reflecting quantity-based processing, working memory load and effortful step-by-step calculation (Delazer et al., 2003). As the responses tended towards retrieval, imaging showed a significant shift of activation to the left angular gyrus, a region of the brain associated with automatic fact retrieval (Delazer et al., 2003; Ung et al., 2020).

This study has several limitations that reduce the strength of its conclusions. Firstly, the sample size was very small (n=13) which limits the extent to which the findings can be applied to broader contexts. Secondly, although the researchers found increased angular gyrus activation, they acknowledged that they could not fully separate direct retrieval from the use of internal procedure. They explicitly note that “the differential effect of strategy use and retrieval cannot be disentangled in the present study” (Delazer et al., 2003, p. 86). Lastly, the neural mechanisms involved in adults learning complex multiplication may differ meaningfully from those involved in children, limiting the direct applicability of the findings to Rocket Math's claim. Despite these limitations, Delazer et al. remains valuable as it is the only study in this report to provide Layer III evidence of the transition from effortful calculation to automatic retrieval.

The relationship between Delazer et al. (2003) and Breneman-Smith (2024) demonstrates Donoghue and Horvath's (2016) concept of upward unpredictability - the principle that findings cannot be reliably translated across non-adjacent layers without investigating the intervening level. Delazer et al. do not attempt to prescribe Layer V classroom pedagogy directly from their Layer III research and Breneman-Smith does not claim to measure specific neural mechanisms.

Conclusion

The effectiveness of the Rocket Math Worksheet Program appears to have a reasonably strong evidentiary basis, although the underlying concept of automaticity is more complex than the program's claims suggest. Evidence from Layer IV provides support for the proposition that rapid responses to number facts are associated with automaticity (Stickney et al., 2012), while Layer V and Layer IV studies provide support for the effectiveness of the program's structured daily peer-practice model (Kohler et al., 1994; Burns et al., 2019; Breneman-Smith, 2024).

Layer III evidence from Delazer et al. (2003) provides a plausible neurological explanation for how repeated retrieval practice may contribute to automaticity, while Díaz-Barriga Yáñez et al. (2023) highlight an important limitation; suggesting that rapid responses may sometimes reflect highly automatised procedures rather than pure memory retrieval.

Overall, the available evidence suggests that Rocket Math's core instructional practices are broadly consistent with the development of automaticity. the available evidence suggests that Rocket Math's core instructional practices are broadly consistent with the development of automaticity.

Whether students achieve this through direct retrieval or highly automatised procedures may be less important than the educational outcome itself. Rocket Math is likely to be an effective approach for developing fluency with basic number facts in primary school students.

References

Australian Curriculum, Assessment and Reporting Authority. (2024a). *Mathematics proficiencies*. https://www.australiancurriculum.edu.au/resources/mathematics-proficiencies/

Breneman-Smith, S. L. (2024). Quasi-Experimental Study of the Impact Spaced Practice and Retrieval Have on Mathematical Fact Fluency in Third-, Fourth-, and Fifth-Grade Students. *Doctoral Dissertations and Projects*.

Cartwright, K. (2022). What does true mathematical fluency look like in the classroom? *Scan: The Journal for Educators*, *41*(6), 8–16.

Crawford, D. (n.d.) *Worksheet Program information*. Rocket Math. https://www.rocketmath.com/worksheet-program-information/

Crawford, D. (2015a, May 30). *What if students become stuck?* Rocket Math. https://www.rocketmath.com/2015/05/30/what-if-students-become-stuck/

Crawford, D. [Rocket Math.com]. (2015b, November 25). *Why teach facts to automaticity?* [YouTube Video]. YouTube. https://www.youtube.com/watch?v=Um7C65AC4Aw

Crawford, D. (2017, March 30). *Do you know the active ingredient in Rocket Math?* Rocket Math. https://www.rocketmath.com/category/why-use-rocket-math/page/2/?srsltid=AfmBOoo_sbzvEwcOsBGUfmsbhnCU1PJsxJU1SY5P6FyToi26u3DbXV-V

Crawford, D. (2020a, March 30). *Change Rocket Math Online Game speed–You can do it, but should you?* Rocket Math. https://www.rocketmath.com/blog/page/10/

Crawford, D. (2020b, February 25). *How Do Online Math Drills Help Children Develop Math Fact Fluency?* Rocket Math. https://www.rocketmath.com/blog/page/10/

Crawford, D. (2022, July 16). *Dr. Don's Worksheet Program Checklist*. Rocket Math. https://www.rocketmath.com/2022/07/16/dr-dons-worksheet-program-checklist/

Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic—An fMRI study. *Cognitive Brain Research, 18*(1), 76–88. https://doi.org/10.1016/j.cogbrainres.2003.09.005

Díaz-Barriga Yáñez, A., Longo, L., Chesnokova, H., Poletti, C., Thevenot, C., & Prado, J. (2023). Neural evidence for procedural automatization during cognitive development: Intraparietal response to changes in very-small addition problem-size increases with age. *Developmental Cognitive Neuroscience, 64*, Article 101310. https://doi.org/10.1016/j.dcn.2023.101310

Dole, S., Carmichael, P., Thiele, C., Simpson, J., O'Toole, C., & Mathematics Education Research Group of Australasia. (2018). *Fluency with number facts—Responding to the Australian Curriculum: Mathematics*. Mathematics Education Research Group of Australasia.

Donoghue, G. M., & Horvath, J. C. (2016). Translating neuroscience, psychology and education: An abstracted conceptual framework for the learning sciences. *Cogent Education, 3*(1), Article 1267422. https://doi.org/10.1080/2331186X.2016.1267422

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539–1552. https://doi.org/10.1037/a0025510

Hasselbring, T. S., Goin, L. I., & Bransford, J. D. (1988). Developing math automaticity in learning handicapped children: The role of computerized drill and practice. *Focus on Exceptional Children, 20(6)*, 1–7. https://doi.org/10.17161/foec.v20i6.7504

Hittmair-Delazer, M., Semenza, C., & Denes, G. (1994). Concepts and facts in calculation. *Brain, 117*(4), 715–728. https://doi.org/10.1093/brain/117.4.715

Kohler, F. W., Ezell, H., Hoel, K., & Strain, P. S. (1994). Supplemental peer practice in a first-grade math class: Effects on teacher behavior and five low achievers' responding and acquisition of content. *The Elementary School Journal, 94*(4), 389–403.

Miller, K. C., Skinner, C. H., Gibby, L., Galyon, C. E., & Meadows-Allen, S. (2011). Evaluating generalization of addition-fact fluency using the taped-problems procedure in a second-grade classroom. *Journal of Behavioral Education, 20*(3), 203–220.

Renaissance Learning. (2003). *MathFacts in a Flash*. Renaissance Learning.

Reynolds, C. R., Voress, J. K., & Kamphaus, R. W. (2015). *Mathematics Fluency and Calculation Test: Technical manual*. PRO-ED.

Rocket Math. (n.d.). *Worksheet Program information*. https://www.rocketmath.com/worksheet-program-information/

Stickney, E. M., Sharp, L. B., & Kenyon, A. S. (2012). Technology-enhanced assessment of math fact automaticity: Patterns of performance for low- and typically achieving students. *Assessment for Effective Intervention, 37*(2), 84–94. https://doi.org/10.1177/1534508411430321

Ung, W. C., Yahya, N., & Tang, T. B. (2020). Detection of mathematical fluency effects on working memory using near infrared spectroscopy. In *2020 IEEE Sensors Applications Symposium (SAS)* (pp. 1–4). IEEE. https://doi.org/10.1109/SAS48726.2020.9220070