Investigating the effect of distributed versus massed tutoring on progress in mathematics

Final Report September 2022

Executive summary

The project

This project investigated whether shorter more frequent sessions of online tutoring have a bigger impact on progress in mathematics attainment than less frequent longer sessions. The project was conducted by online tutoring company, PLYTIME Learning, and Dr Steph Ainsworth from the Education and Social Research Institute (ESRI) at Manchester Metropolitan University (MMU). 130 children, aged 8 to 10 years, were randomly assigned to 3 groups: a distributed practice group receiving 3 x 15 minutes of online FOCUS tutoring each week; a massed practice group receiving 1 x 45 minute of online FOCUS tutoring each week; and a control group who received no tutoring and had no access to the platform. Students in the first two groups received 6 weeks of tutoring. A measure of mathematics attainment was taken before and after the tutoring period. A measure of mathematics progress was also taken.

Key conclusions
Children who received FOCUS tutoring scored significantly higher on mathematics attainment than children in the control group. The result was statistically significant and suggests that FOCUS tutoring improves attainment.
Children who received 3 x 15-minute tutoring sessions scored significantly higher on a measure of mathematics progress than students in the 1 x 45-minute group. The result was statistically significant and suggests that shorter, frequent tutoring is more effective than longer weekly sessions.
Descriptively the same pattern of results was found for the measure of mathematics attainment (rather than progress); children who received shorter more frequent tutoring sessions scored higher on average than those receiving one longer session a week, but the difference between the distributed and massed practice groups fell just short of the threshold for statistical significance (we can say with 93% confidence that the observed difference in scores is due to the difference in tutoring conditions rather than due to chance, whereas the threshold for statistical significance is 95%).

Overall, the results suggest that shorter more frequent online tutoring sessions are more effective than longer weekly sessions in supporting progress in mathematics in children aged 8 to 10 years.

1. Introduction

Background

The question of whether mathematics practice is more effective when conducted in a massed (all at once) or distributed (spread out over shorter sessions) form has been the subject of previous research (e.g. Carpenter et al., 2012; Codding et al., 2016; Schutte et al. 2015). Overall, the results suggest that distributed practice is more effective, however the effect size varies depending on a number of factors, e.g. type of practice/instructional input, length of time of each session, spacing of each session, online versus face-to-face instruction, and the measures used to assess attainment (Schutte et al., 2015; Codding et al., 2016; Leung et al., 2005). While the relative benefits of massed versus distributed practice are theoretically interesting, they are also of commercial interest to tutoring companies, who are aiming to provide cost-effective and time efficient tutoring for students which leads to a significant impact on attainment. Within this project, tutoring company, PLYTIME Learning, were interested in exploring whether their students would experience greater gains in mathematics attainment when they experienced three short tutoring sessions of 15 minutes (their FOCUS 15®progamme) versus one longer session of 45 minutes each week. The research question and objectives for the project are as follows.

Research question: Do children experience greater gains in mathematics attainment when they engage with a distributed package of tutoring (3x15 minutes a week) versus a massed package of tutoring (1x45 minutes a week)?

Objectives:

- To evaluate the difference in mathematics attainment between children who have engaged with distributed versus massed tutoring, while taking into account prior attainment.
- To evaluate the difference in both the massed and distributed groups' mathematics attainment relative to a control group, who have not received any tutoring.
- To evaluate the difference in mathematics progress between children who have engaged with distributed versus massed tutoring.

2. Methods

Project Design

The project ran from May to August 2022. PLYTIME recruited 130 students who were currently in year 4 or year 5 of primary school (in England this corresponds to ages 8 to 10 years). This age group was chosen as the focus of the study as there is a high demand for online tutoring as children prepare for their year 6 SATS (Standard Assessment Tests) and in

some cases high school entrance exams. The children were randomly assigned to one of three experimental groups:

- **Group D: the distributed practice group**, which received 3 x 15 minutes a week of online FOCUS tutoring for 6 weeks.

- **Group M: the massed practice group**, which received 1 x 45 minutes a week of online FOCUS tutoring for 6 weeks.

- **Group C: the control group**, which did not receive any tutoring or have access to the platform during the project, but were given a voucher providing them with the opportunity to redeem an equivalent value of tutoring at the end of the project.

There were 45 children each in Group D and Group M and 40 children in Group C. Students in Groups D and M who received FOCUS Tutoring had unlimited access to the PLYTIME Learning platform to practice as directed by their tutors. Students in Groups C did not have access to the platform until after the project.

All tutoring sessions were one-to-one and took place online via PLYTIME's online system using the Twilio video conferencing platform. All tutors had enhanced DBS clearance (from the Disclosure and Barring Service). Each tutoring session had the same structure and followed the same pedagogical approach. The tutor and student would review questions that the student had answered incorrectly or flagged for support, with tutors using a variety of methods to explain how to reach the correct answer. If the student was consistently struggling with a particular skill the tutor would then create a 'Mix' – an algorithm that limits questions to a chosen theme, topic or skill for the student to focus practice on before the next session. This scaffolded approach, is designed to support students to master skills as they progress through the curriculum onto more advanced topics.

Recruitment and ethical considerations

The students were recruited by PLYTIME Learning. Participant information sheets and consent forms were sent out to local schools to be shared with parents. They were also shared with existing parental contacts and through social media. Written consent was gathered from parents, and assent from children was gathered via an assent form. Care was also taken to monitor children's engagement through the project to ensure that they remained happy to continue with the project throughout. Participants were offered an incentive to take part. All parents of children who completed the full 6 weeks of tutoring sessions as well as the pre- and post- tests were reimbursed for the initial 'buy-in' amount of £29. The children also received a thank you voucher (Amazon) for £20 (passed on via the parents). It was made clear to both parents and children that they were able to withdraw from the project at any time without giving a reason. The project was approved by the Faculty of Health and Education Ethics Committee at Manchester Metropolitan University.

Measures

A measure of mathematics attainment was taken both at the beginning of the project (before students received any tutoring) and at the end after they had had their 6 weeks of

tutoring sessions (or after 6 weeks of time had passed for the control group). For some students the post-tutoring test occurred after 7 weeks rather than 6. This was to allow flexibility in terms of students being on holiday or ill and needing to delay either a week of tutoring or the final test until the following week. A measure of mathematics progress was also taken at the end of the tutoring period. Details of the measures used are provided below.

1. **Fixed measure of mathematics attainment** – this measure developed by PLYTIME and reviewed by Steph Ainsworth (MMU) was designed to assess students' attainment in the following areas of mathematics: addition, subtraction, multiplication, division and fractions (including decimals). Within this test, questions based on the content of the Year 4 and Year 5 National Curriculum were included. Children were given 20 minutes to complete the test online. There were two different versions of the test, designed to be of equivalent difficulty and with equivalent coverage of the curriculum, with 50% of students taking Test 1 at pre-test and 50% taking Test 2. At post-test, children who took Test 1 at pre-test took Test 2 at post-test and vice-versa. This was done to mitigate for the possibility that one of the tests might happen to be slightly more difficult than the other. All participating students took this test at the beginning and end of the project (130 children).

2. **Adaptive measure of progress in mathematics attainment** – While the *fixed measure of mathematics attainment* (described above) is fixed in the sense that students of all abilities were given the same level of test, this second measure was adaptive. In other words, the questions that students were asked to complete depended on their current level of attainment. This measure was generated by PLYTIME's digital tutoring system in response to students' performance as they engaged with the practice activities built into the PLYTIME Learning system. Points were awarded to the student based on question difficulty and time taken to answer the question within an allotted time. This measure was collected for FOCUS Tutoring Groups D and M only (90 children) as Group C did not have access to the platform.

   Within the PLYTIME tutoring system, the mathematics curriculum that the students engage with is split into themes, topics and endcodes. Endcodes are the lowest level of detail when chunking the curriculum, with each one representing a distinct skill, e.g. the encode 'Multiply 1 digit number by a 3-digit number', sits within the topic of 'Multiplication Arithmetic', which in turn is part of the theme 'Numeracy Arithmetic'. As students answer questions within the system, they build a performance score at an endcode level based on whether they answered the question correctly and how long it took to answer. If a student is struggling, the endcode score will decrease to a level where easier questions may be available; conversely, as a student improves, their endcode score will increase, allowing them to tackle more challenging questions. The *adaptive measure of progress in mathematics attainment* used in the analysis is the net total of the changes in endcode performance for all the endcodes attempted over the tutoring period. This provides an overall measure of a student's progress across the different skills.

3. Results

Outcomes and analysis

Statistical analyses were conducted on the data to investigate whether there were significant differences in the scores between the three groups (D, M and C) on students' mathematics attainment and progress.

1) Fixed measure of mathematics attainment

Table 1 shows the mean scores for each group on the *fixed measure of mathematics attainment*. The standard deviation (a measure of how much students' scores varied around the mean) and the number of children in each group is also shown.

Table 1: Fixed measure of mathematics attainment for each of the three groups

Group	Mean	Standard deviation	Number of children
Group D	28.24	8.30	45
Group M	27.36	7.51	45
Group C	22.63	7.87	40

One-way analysis of covariance (ANCOVA) was conducted on the data to see if there was a significant effect of group (D, M or C) on mathematics attainment. The students' baseline scores (taken before the tutoring took place) were included as a covariate to account for the fact that we would expect children with relatively high scores at baseline to also have relatively high scores after the tutoring.

As expected, the covariate, pre-tutoring score on the fixed measure of mathematics attainment, was significantly related to the post-tutoring score, $F(1,126) = 390.31$, $p<.001$, $\eta 2 = .76$. The ANCOVA revealed a significant effect of experimental group on the *fixed measure of mathematics attainment*, $F(2,126) = 34.65$, $p<.001$, partial $\eta 2 = .36$. In other words, there was an overall difference in scores depending on which group children were in.

In order to investigate if there was a significant difference between the individual groups, planned contrasts were conducted. The mean score for group M was found to be significantly higher than for group C, $t(83)=5.51$, $p<.001$ (one-tailed). That is, children who received either the distributed or massed FOCUS tutoring sessions performed significantly better than those children in the control group who received no tutoring and had no access to the platform.

While Table 1 shows that the children in the distributed practice group scored higher than the children in the massed practice group, this difference did not quite reach the threshold for statistical significance (the standard threshold set for statistical significance is $p=0.05$), $t(88)=1.22$, $p=.072$ (one-tailed). While Group D scored higher than Group M on average, the significance is marginal, meaning this analysis doesn't conclusively rule out the possibility that this difference could be due to chance rather than due to the difference in the different tutoring conditions. In other words, a difference as large or larger than the one found between groups D and M would occur 7.2% of the time by chance even if there were no

true difference between the treatments, which is just outside the threshold for statistical significance of 5%. It is possible that this marginal result may have reached significance if some of the limitations outlined towards the end of the report had been eased. Therefore, while the results suggest that attainment was higher for Group D than Group M, further research using a larger sample would be needed to provide more conclusive evidence.

2) Adaptive measure of progress in mathematics attainment

While the analysis above investigated differences in attainment (differences between the overall test scores after 6 weeks), this second analysis explored differences in mathematics progress from the beginning to the end of the tutoring period. This analysis only included the two groups that received tutoring (Group D and Group M). Table 1 shows the means and standard deviations for the *adaptive measure of progress in mathematics attainment.*

Table 2: Adaptive measure of progress in mathematics attainment

Group	Mean	Standard deviation	Number of children
Group D	8.77	18.67	45
Group M	3.23	8.58	45

A t-test was conducted to investigate if the difference between groups D and M was significant. The test showed that the mean progress in mathematics attainment for the students in the distributed practice group was significantly higher than the mean progress for the students in the massed practice group, t(88)=1.81, p=.037 (one-sided) and passed the threshold for statistical significance. This difference reflects a medium to large effect (Glass's $\Delta = 0.65$) of distributed versus massed tutoring on progress (Glass's Δ is reported, rather than Cohen's d to account for the difference in standard deviations between the groups (Glass, 1981)).

5. Conclusion

Key conclusions
Children who received FOCUS tutoring scored significantly higher on mathematics attainment than children in the control group. The result was statistically significant and suggests that FOCUS tutoring improves attainment.
Children who received 3 x 15-minute tutoring sessions scored significantly higher on a measure of mathematics progress than students in the 1 x 45-minute group. The result was statistically significant and suggests that shorter, frequent tutoring is more effective than longer weekly sessions.
Descriptively the same pattern of results was found for the measure of mathematics attainment (rather than progress); children who received shorter more frequent tutoring sessions scored higher on average than those receiving one longer session a week, but the difference between the distributed and massed practice groups fell just short of the threshold for statistical significance (we can say with 93% confidence that the observed difference in scores is due to the difference in tutoring conditions rather than due to chance, whereas the threshold for statistical significance is 95%).

Overall, the results suggest that shorter more frequent online tutoring sessions are more effective than longer weekly sessions in supporting progress in mathematics in children aged 8 to 10 years.

Interpretation

The fact that both tutoring conditions led to substantial gains in attainment in comparison to the control group provides evidence for PLYTIME Learning's FOCUS Tutoring approach, which is adapted to the students' needs, allowing them to work through the curriculum at an appropriate pace (c.f. Fuchs et al., 2008; Hasselbring et al., 1988; Martens and Eckert, 2007). The results suggest that a distributed approach to online mathematics tutoring, where children receive shorter more frequent sessions, may be more effective at promoting progress than an approach where students receive less frequent longer sessions. The findings of this study align with those of Codding et al. (2016), who investigated the impact of frequency of mathematics instruction in the context of a small-group face-to-face intervention. They also found that shorter sessions conducted several times a week were more effective than one longer session a week (controlling for total duration). These findings are in line with the broader literature which has demonstrated the benefits of shorter more frequent learning opportunities across a range of subjects/tasks (e.g. Carpenter et al., 2012; Cepeda et al. 2006; Cepeda et al., 2011; Codding et al., 2011; Martens et al., 2007).

Intuitively we might expect children to find it easier to concentrate during shorter sessions, especially given that their capacity for sustained attention is continuing to develop within the age range studied (Betts et al., 2006). The findings may also be interpreted in the context of cognitive load theory (Sweller, 1988), which highlights the limited nature of working memory and people's capacity for information processing. According to this theory, learning may be maximised by using pedagogical approaches which aim to minimise cognitive overload and optimise use of working memory (Baker, 2022). A distributed approach where practice is spaced out over time rather than completed in one go, has been argued to support remembering through more frequent retrieval of memory schemas, which in turn strengthens the links between the different components of the schema. In addition the modelling of the problems and scaffolding provided by the tutors within the FOCUS tutoring sessions, may have further supported the students in developing established schemas without their working memory being overloaded (Rawson et al., 2011), and the frequent formative assessments that took place throughout the tutoring (as students engaged with the practice questions within the PLYTIME system), provided regular opportunities for students to retrieve and apply their developing knowledge (Roediger et al., 2011).

In summary, this study adds to the evidence base in support of distributed versus massed practice by demonstrating this effect within the context of online mathematics tutoring for children aged 8 to 10 years. The findings of this study may be of interest to tutors who are designing the structure of their programmes and to parents who are choosing models of instruction for their children.

Limitations

The sample of students involved in this project was relatively modest (n=130), which may explain why the difference between the two tutoring conditions failed to reach the threshold for statistical significance for the *fixed measure of mathematics attainment*. It is possible that the different tutoring conditions would have revealed a significant effect for the fixed measure if a larger sample had been used, given the closeness of the p value to the threshold (p=.074). Another factor that might have limited the study's power to detect differences between the groups is the fact that students were only given 20 minutes to complete the test, which could have limited their scoring capability. On average, students in Group D scored 21.78 in their first test and 28.24 in their second test, an increase of 6.47 or 30%, and students in Group M scored 22.13 in their first test and 27.36 in their second test, an increase of 5.22 or 24%, and students in Group C scored 22.98 in their first test and 22.63 in their second test, which shows no increase across the period. Whilst the results for FOCUS Tutoring groups do show significant improvements in attainment across the period overall, it is possible that a longer testing time would have allowed the students to have completed a greater number of questions (with a maximum of 80 points available), potentially allowing greater differences between the distributed and massed conditions to have been observed.

It is also important to note that students who signed up to FOCUS tutoring were not specifically practicing question types related to the fixed measure of attainment. Instead, students in Groups D and M who signed up to FOCUS tutoring, practiced questions personalised to their level of ability in Maths which may from time to time have included similar questions as contained in this measure. In other words, they were not practicing questions solely to improve the results of the measure and some students may have practiced similar questions as contained in the measure more than others. This would suggest there may be transferability from FOCUS tutoring to more distal outcomes. However, use of a standardised measure of mathematics would be needed to test this more conclusively.

References

Baker, L. A. (2022). The utility of distributed practice in curriculum design and effective learning strategies. *Impact, 14*.

Betts, J., Mckay, J., Maruff, P., & Anderson, V. (2006). The development of sustained attention in children: the effect of age and task load. *Child Neuropsychology, 12*(3), 205-221.

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., and Pashler, H. (2012). Using spacing to enhance diverse forms of learning: review of recent research and implications for instruction. *Educational Psychology Review, 24*(3), 369–378.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380.

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., and Pashler, H. (2009). Optimizing distributed practice. *Experimental Psychology, 56*(4), 236–246.

Codding, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency interventions: A component analysis. *Learning Disabilities Research & Practice*, *26*(1), 36– 47.

Codding, R. S., VanDerHeyden, A. M., Martin, R. J., Desai, S., Allard, N, & Perrault, L. (2016). Manipulating Treatment Dose: Evaluating the Frequency of a Small Group Intervention Targeting Whole Number Operations. *Learning Disabilities Research & Practice, 31*(4), 208-220.

Cole, R., Haimson, J., Perez-Johnson, I., & May, H. (2011). Variability in Pretest-Posttest Correlation Coefficients by Student Achievement Level. Research Report. Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences.

Fuchs, L. S., Fuchs, D., Powell, S. R., Seethaler, P. M., Cirino, P. T., & Fletcher, J. M. (2008). Intensive intervention for students with mathematics disabilities: Seven principles for effective practice. *Learning Disability Quarterly, 31*(2), 79–92.

Glass G. V., McGaw B., Smith M. L. (1981). *Meta-Analysis in Social Research*. Beverly Hills, CA: Sage.

Martens, B., & Eckert, T. (2007). The Instructional Hierarchy as a Model of Stimulus Control over Student and Teacher Behavior: We're Close but are We Close Enough? *Journal of Behavioral Education*, *16*(1), 82–90.

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.

Martens, B. K., Eckert, T. L., Begeny, J. C., Lewandowski, L. J., DiGennaro, F. D., Montarello, S. A., et al. (2007). Effects of a fluency-building program on the reading performance of low-achieving second and third grade students. *Journal of Behavioral Education, 16*(1), 39–54.

Leung, K. C. (2018). An updated meta-analysis on the effect of peer tutoring on tutors' achievement. *School Psychology International, 40*(2), 200-214.

Rawson, K. A., Thomas, R. C. and Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review*, *27*(3): 483–504.

Roediger, H. L., Putnam, A. L., & Smith, M. (2011). Ten Benefits of testing and their applications to educational practice. *In* Mestre JP and Ross BH (eds) *Psychology of Learning and Motivation*. Academic Press, pp. 1–36.

Schutte, G. M., Duhon, G.J., Solomon, B. G., Poncy, B. C., Moore, K., & Story, B. (2015). A comparative analysis of massed vs. distributed practice on basic math fact fluency growth rates. *Journal of School Psychology, 53*(2), 149-159.

Sweller J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science, 12*(2), 257–85.