Times Tables Rock Stars: An Academic Critique

Julia Thomson and Sarah Wright

Abstract

Times Tables Rock Stars has become a popular digital intervention used in schools across the world. Through a mobile phone and tablet app schools can set homework and competitions in their school to encourage children to practice their times tables both during and outside of school. The app uses colourful characters and gamified incentives, such as the opportunity to earn coins, to engage pupils and keep them motivated in their times tables practice. In England Times Tables Rock Stars is marketed to primary schools in the context of the now statutory multiplication tables check (MTC) for pupils at the end of year four. Schools who sign up are able to access specific support in relation to the MTC, including a game within the app which mirrors the format of the MTC test. However, as yet there have not been any formal research studies exploring the effectiveness of Times Tables Rock Stars. Therefore, the following critique explores other research into similar digital interventions to start to address the question of whether Times Tables Rock Stars could be a useful intervention for increasing pupils' multiplication tables fluency.

Times Tables Rock Stars: An Academic Critique

In 2018 a multiplication tables check (MTC) for pupils at the end of year four was introduced (Standards and Testing Agency, 2022), becoming a statutory requirement in 2021. The purpose of the assessment is to enable teachers to check which pupils may need help with reaching the goal of knowing all multiplications up to 12, "off by heart" by the end of Key Stage Two (Gibb & Department for Education, 2018). This goal is linked to the principle of the maths mastery pedagogy, which, with links to cognitive load theory (Sweller, 1988), suggests that learning basic number facts, such as multiplication tables, by rote, frees up working memory, enabling focus on more complex problem solving (NCETM, 2017).

The MTC assessment is delivered online with pupils given six seconds to enter their answer (Standards and Testing Agency, 2022). This may mean teachers are keen to find ways to familiarise pupils with a digital format before the MTC takes place. In addition the experience of sudden school closures due to COVID-19 (Hume et al., 2023) has led school leaders to look to increase technology use in schools to support a more strategic approach to any future needs for remote learning (Floyd et al., 2023). The digital multiplication tables fluency intervention (DMFI), Times Tables Rock Stars (TTRS; Reddy, 2010) may therefore appeal to school staff as it uses technology to support pupils to increase their multiplication table fluency.

What is Times Tables Rock Stars?

The Times Tables Rock Stars website reports that over 17,000 schools are signed up worldwide (Maths Circle Ltd., 2024) making it perhaps one of the most popular digital interventions used in schools. Created in 2010 by Bruno Reddy (Ward, 2015), TTRS has developed from an impromptu classroom game into a digital app available on multiple devices worldwide (Maths Circle Ltd., 2024). TTRS costs between £105.85 and £178.85 per year for a school subscription (Maths Circle Ltd., 2024) and includes access to pupil progress data. Schools are encouraged to celebrate success, for which printable materials are provided alongside unlimited access to the DMFI.

TTRS is designed for pupils aged six and over as a whole class or targeted intervention for those identified as needing specific support with multiplication tables. Once signed up pupils select a 'Rock Name' and design their avatar, which can then be further accessorised as 'coins' are earnt through game play. There are eight games in TTRS, three in multiplayer mode and five in single player mode. In multiplayer mode pupils compete anonymously against all TTRS users worldwide or within their school. Each multiplayer game follows the same format with one minute to answer as many multiplication questions as possible correctly. Single player mode includes multiple games with varying levels of adaptability to pupils' current level of fluency. Many games also include an element of progress tracking available to pupils, enabling them to view their progress over time. TTRS also includes specific support for school staff around the MTC (Maths Circle Ltd., 2024), with the game 'Sound Check' designed to mimic the MTC.

Pupils complete TTRS independently, but school staff can set homework and competitions called 'battles' where pupils in teams compete to produce the average highest number of correct answers per team member. Pupils can view their current progress against peers by looking at class and school wide ranking tables, with only pseudonyms visible to pupils.

Theoretical Underpinnings

Although information about TTRS states that it is "pedagogically sound" (Maths Circle Ltd., 2024), there is no explicitly stated theoretical basis to justify its design. However, from a psychological perspective, explicit links to the MTC and focus on increased speed of recall could link to the theoretical area of fluency. Furthermore, gamified incentives in the design of the TTRS app and encouragement for schools to celebrate success, suggest the influence of psychological theories related to motivation.

Fluency

The concept of fluency can be traced back to Haring and Eaton's (1978) instructional hierarchy. This theory of learning suggests that learners must pass through a series of four sequential steps in order to learn a new skill. In this model fluency is the second step after initial ‘acquisition’ and is needed for learners to reach a level of proficiency that enables them to generalise and adapt skills they have learnt to novel situations (Haring & Eaton, 1978). Fluency measurement depends on the context of the individual skill (Haring & Eaton, 1978). When it comes to multiplication tables, fluency is often synonymous with automatic recall, or ‘automaticity’ (Carr & Alexeev, 2011; Hasselbring et al., 1987; Woodward, 2006). Early arithmetic fluency has been found to correlate to higher mathematics attainment later on (Carr & Alexeev, 2011). One explanation for this is that automaticity of basic number facts enables the learner more cognitive capacity to focus on higher mathematical concepts (Hasselbring et al., 1987). The goal of most TTRS games is to increase the speed at which answers are given and under fluency in the pupil’s stats section of the app it states “Under 3s/q is considered to be automatic recall.”(Reddy, 2010). This links to Hasselbring and colleagues (1987) conceptualisation of automatic recall as giving correct answers in one or two seconds.

Fluency can be increased through practice, repetition and over learning (Haring & Eaton, 1978). Research into increasing multiplication fluency has found that repeated timed practice, often referred to as ‘drill and practice’ is an effective method (Codding et al., 2011; Woodward, 2006). This again links to the design of TTRS games, the majority of which are completed under a set time limit, with pupils being encouraged to complete multiple games repeatedly in each session. Furthermore, additional research into digital maths games has found that they are most effective at increasing fluency when immediate feedback is given (Hawkins et al., 2017; Moyer-Packenham et al., 2019), which is another key feature of TTRS.

Motivation

Motivation is the process by which individuals initially and continuously take part in activities, often linked to wider goals (Schunk & DiBenedetto, 2020). Although the key mechanism through which fluency is achieved is repetition, this may not be interesting enough to engage learners at the level necessary for the required outcome. Therefore, Haring and Eaton (1978) suggest that reinforcement may be needed to provide “the necessary motivation to continue practising" (Haring & Eaton, 1978, p. 28). This is in line with Bandura's (1986) social cognitive theory view of motivation and learning, which suggests that internal motivation is only possible with initial external support. Bandura (1986, p. 240) states "without the aid of positive incentives during early phases of skill acquisition, potentialities are likely to remain undeveloped." Meaning that those at an early stage of developing fluency will need external support to motivate them to put in the practice necessary to reach the required level of fluency to fully embed the skill. Digital games have been found to increase motivation in comparison to traditional teaching methods (Fadda et al., 2022), with in game immediate rewards being an important part of this (Hawkins et al., 2017). As well as encouraging school staff to celebrate success, coins received for correct answers which can be used to 'buy' digital accessories for pupil's avatars could also be viewed as a TTRS design feature linked to increasing motivation to practice multiplication tables.

Although encouraging external motivators at the initial stage, Bandura's (1986) social cognitive theory's view of motivation largely focuses on ways in which an individual is internally motivated. A key aspect of this is self-efficacy, which is the idea that an individual must believe they are able to achieve a goal through their own actions in order to be motivated to complete the steps required to meet that goal (Bandura, 1986). Research has shown that self-efficacy is a key mediator in the positive link between motivation and longer term outcomes when it comes to mathematical attainment (Gottfried et al., 2013; Skaalvik et al., 2015). TTRS may encourage increased sense of self-efficacy through the ability of pupils to look back at their achievements over time, which is a key feature of the 'stats' section of the TTRS app.

Impact and Effectiveness of Times Tables Rock Stars

There have not yet been any research studies exploring the impact of TTRS as an intervention for increasing multiplication tables fluency. The apparent popularity of the intervention may therefore be linked to school staff being more likely to choose interventions through peer recommendation than research evidence and word of mouth has led to the large take up (Pegram et al., 2022). However, despite the powerful impact of peer recommendation, it is important for educators to have access to a research evidence base when choosing interventions. Previous literature reviews and meta-analysis have found digital interventions to be a successful method of increasing mathematical fluency in general (Codding et al., 2011; Cozad & Riccomini, 2016) but to date there have been no reviews focusing specifically on multiplication fluency. Therefore, a literature review was conducted to answer the question: 'Are digital games effective in increasing multiplication fact fluency in school aged children and young people?' Details of the search strategy, specific inclusion criteria and search terms used can be found in Appendix A. A total of 48 papers were initially found, with 42 papers removed after screening, leaving a total of six papers included in the following literature review. All included papers were quality assessed using an adapted version of the Downs and Black (1998) checklist (Appendix B) and a data extraction table was created (Appendix C) to enable synthesis in the evaluation.

Overview of Included Studies

Alongside the search of academic journals, a grey literature search was conducted to allow for a broader picture of the evidence (Paez, 2017). Therefore, two of the six included papers are doctoral theses (Agee, 2019; Smith, 2010). Overall, included studies span a time frame of ten years, with the oldest study published in 2010 (Smith, 2010) and the most recent in 2023 (Kromminga & Codding, 2023). All studies took place in school environments, although most took place in only one or two schools meaning results may not generalise to other settings. Participants were pupils aged between seven and fourteen years with overall numbers ranging from 29 to 166. Only one study is from the UK (Jay et al., 2019), with the rest taking place in the USA (Agee, 2019; Berrett & Carter, 2018; Denham, 2013; Kromminga & Codding, 2023; Smith, 2010). Smith (2010) used a screening multiplication assessment to select eligible participants, with the inclusion criteria of all pupils scoring fewer than 80 correct answers in two minutes, which overlaps with Hasselbring and colleagues (1987) definition of automatic recall, which would equate to between 60 and 120 correct answers in two minutes..

There are several different terms used to identify pupils in need of additional support in relation to disability or learning needs. In the United States of America where most of the included studies in this critique took place, the term ‘registered disability’ is often used. However, as this critique is related to the national curriculum in England, the term special educational needs (SEN) will be used to refer to these pupils. Denham (2013) and Jay and colleagues (2019) explicitly excluded pupils with SEN in their studies. Berrett and Carter (2018) noted that 14% of students attending the school at which their study took place, had SEN, but did not specify the number in relation to their participants. The remaining three studies (Agee, 2019; Denham, 2013; Kromminga & Codding, 2023) specify that between 12% and 17% of their participants had known special educational needs.

Although no studies met the quality definition of a randomised controlled trial (RCT; Connolly et al., 2018), four of the studies did include a control comparison group. Berrett and Carter (2018) implemented a multiple baseline design with all groups baseline period including playing a literacy digital game. In all other studies where a control group was used it was unclear what ‘business as usual’ meant for these pupils and for Agee (2019) the control group used had a significantly higher level of fluency at baseline than one intervention group, making the comparison less robust. Two studies also included a peer tutoring intervention comparison group (Kromminga & Codding, 2023; Smith, 2010) and two included comparisons with altered delivery of the same intervention (Agee, 2019; Denham, 2013).

Two studies used DMFIs with digital flash-cards (Kromminga & Codding, 2023; Smith, 2010) and the others used DMFIs incorporating interactive games including some form of token reward for correct answers (Agee, 2019; Berrett & Carter, 2018; Denham, 2013; Jay et al., 2019). Most interventions were implemented for short periods daily and continued for two to seven weeks, to follow a school term, or in line with the resources available to the research team. However, Denham (2013) completed their intervention over only two days with 30 minutes game play on each, meaning the results may not fully represent the potential impact of the intervention.

Measures used in all studies included a paper and pencil multiplication fact fluency assessment, with most recording results in terms of number of correct answers in one or two minutes. Two studies also used measures to gather pupil views on the intervention (Denham, 2013; Kromminga & Codding, 2023) and Agee (2019) included a measure of pupil's mathematical self-efficacy.

Results of Included Studies

All studies reported improvements in rates of multiplication fluency from pre- to post-test, however there were varying degrees of significance between the improvements of control and intervention groups. Results for the two studies using digital flash card interventions were limited. Kromminga and Codding (2023) found no significant difference between groups, with the exception of a significantly higher level of progress in the digital only intervention group than the group who received a mixture of digital and peer tutoring interventions. However, this may be an anomaly in the group rather than an impact of the intervention as there was no significant difference between the group that received only peer intervention and either the digital only or combined group. Furthermore, the group sizes in this study mean that the statistical analysis was likely to be under powered, meaning the analysis cannot be fully relied upon. An interesting descriptive statistic finding in this study however was that the digital only intervention made more progress in the digitally delivered fluency measure than either of the comparison groups, which may point to some rehearsal benefit of mode of assessment. However, again the relatively small group size and lack of statistical significance testing means this finding must be viewed with caution.

When it came to the more interactive DMFIs Jay and colleagues (2019) found a significant difference between improvements in arithmetic fluency between control and intervention groups at the mid-point of their crossover design, and again from mid- to post-test. However, their participant group did not include pupils with learning difficulties and the school from which the sample was taken was described as having 'average' attainment levels in mathematics. Therefore, these results may not compare to pupils who are in most need of support to increase their multiplication fluency.

Furthermore, the intervention enabled pupils to select between number bond and multiplication questions, with results showing a correlation between time spent on multiplication questions and higher increases in multiplication fluency. Although this links to the need for high levels of repetition to achieve fluency (Haring & Eaton, 1978), it is not clear if the pupils who completed the most multiplication questions were also those with the highest level of self-efficacy at the start. This may have impacted upon pupils' motivation to complete more multiplication tasks (Bandura, 1986) rather than the game itself having a specific impact upon this. This is of particular note when viewed in the context of results from Berrett and Carter (2018), who found a correlation between prior ability level and gains made through the intervention, with those with the lowest level of mathematical attainment making the lowest level of gains. However, this study also recorded a strong effect size for the interaction between baseline and intervention periods for all groups, indicating the intervention resulted in improvements in multiplication fluency when compared to no intervention. In addition Agee (2019) also found a significant increase in mathematical self-efficacy from pre- to post-test for participants in the intervention group, meaning that it may be possible for DMFIs to increase all learners self-efficacy and therefore increase their motivation.

Conclusions and Limitations of Literature Review

While all studies included in this literature review did find some improvements in multiplication fluency from pre- to post-intervention, the majority also found improvements in comparison and control groups. Furthermore, the quality of studies in general was low with no RCTs and many including low participant numbers meaning statistical analysis may be unreliable. Furthermore, Berrett and Carter (2018) was the only study to take into account missing data by replacing with average scores from other measures. In all other studies participants with missing data were excluded, meaning results may not translate to real world impact as they do not take into account pupils who may have difficulty accessing all aspects of the intervention. This matters because interventions may be used by schools to target pupils who are falling behind expected levels of progress, including those with SEN and so it is important that research takes the accessibility of interventions to all pupils into account.

There were flaws in the included studies, but they all pointed towards the same positive impact of DMFI on multiplication fluency. In addition, the fact that there were more significant results for interventions taking the form of interactive games highlights the potential positive role incorporating motivational elements into DMFI may have on their effectiveness.

Implications for Practice and Recommendations for Future Research

The literature review did not highlight any firm evidence to support the effectiveness of TTRS. However, overall DMFIs were found to increase multiplication fluency particularly when involving gamification and when played consistently over several weeks. Therefore, when educators make choices around multiplication fluency interventions, or when Educational Psychologists (EPs) are advising on them, the most important things to consider are the specific motivational features of the intervention and time needed to support pupils to access the intervention on a regular basis. However, caution should be taken when using such interventions in a targeted way as the current limited evidence focuses predominantly on using such interventions at the whole class level. Results that do include learners with a higher level of need indicate that DMFIs have less impact for this group than they do for those with existing higher levels of mathematical attainment (Berrett & Carter, 2018). This will have particular relevance to those working in specialist settings as although TTRS contains accessibility features, the fact that DMFIs appear to have less impact for pupils with SEN may mean educators and EPs would be better placed to explore in-person interventions instead. Future research would benefit from exploring the potential impact of DMFIs and specifically TTRS on pupils with SEN and how these interventions may be adapted to suite all learners.

In relation to the specific context of some school leaders potentially choosing TTRS to support pupils to prepare for the MTC, there was again limited evidence from the reviewed studies to support DMFIs over traditional teaching methods. However, the finding from Kromminga and Codding (2023) that the digital intervention only group made more progress in the digitally administered multiplication fluency measure, might cautiously be taken to indicate some benefit for pupils having the opportunity to practice the method of assessment delivery. Future studies are needed to fully explore the impact of rehearsal effects on test performance in the context of digitally delivered interventions in comparison with more traditional methods.

Although research evidence is important in enabling educators to make well informed decisions when choosing interventions (Connolly et al., 2018; Pegram et al., 2022), it can never replace the need for monitoring pupil progress in order to assess if the intervention is right for individual settings. In this way TTRS may prove to be most beneficial due to its in-built data collection which will enable school staff to see for themselves if it works for their school. This feature could also be utilised by future researchers to gather data on the broader effectiveness of TTRS as an overall intervention as well as for specific pupil groups.

References

Agee, F. R. (2019). *Effect of Technology-assisted Learning Platform on Multiplication Fact Fluency and Self-efficacy* [Ed.D.Ed.Lead., Union University]. https://www.proquest.com/docview/2290955610/abstract/1E6F59FB307649D7PQ/1

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Berrett, A. N., & Carter, N. J. (2018). Imagine Math Facts Improves Multiplication Fact Fluency in Third-Grade Students. *Journal of Behavioral Education*, *27*(2), 223–239. https://doi.org/10.1007/s10864-017-9288-1

Carr, M., & Alexeev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of arithmetic strategies. *Journal of Educational Psychology*, *103*(3), 617–631. https://doi.org/10.1037/a0023864

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. https://doi.org/10.1111/j.1540-5826.2010.00323.x

Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016. *Educational Research*, *60*(3), 276–291. https://doi.org/10.1080/00131881.2018.1493353

Cozad, L. E., & Riccomini, P. J. (2016). Effects of Digital-Based Math Fluency Interventions on Learners with Math Difficulties: A Review of the Literature. *Journal of Special Education Apprenticeship*, *5*(2). https://eric.ed.gov/?id=EJ1127743

Denham, A. R. (2013). Strategy Instruction and Maintenance of Basic Multiplication Facts through Digital Game Play. *International Journal of Game-Based Learning (IJGBL)*, *3*(2), 36–54. https://doi.org/10.4018/ijgbl.2013040103

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, *52*(6), 377–384. https://doi.org/10.1136/jech.52.6.377

Fadda, D., Pellegrini, M., Vivanet, G., & Zandonella Callegher, C. (2022). Effects of digital games on student motivation in mathematics: A meta-analysis in K-12. *Journal of Computer Assisted Learning*, *38*(1), 304–325. https://doi.org/10.1111/jcal.12618

Floyd, A., Baxter, J., Morales, A., & Bari, R. (2023). Leading online learning during a pandemic and beyond: Challenges and opportunities for school leaders in England. *Educational Management Administration & Leadership*. https://doi.org/10.1177/17411432231191171

Gibb, Rt. H. N., & Department for Education. (2018). *Multiplication tables check trials to begin in schools*. GOV.UK. www.gov.uk/government/news/multiplication-tables-check-trials-to-begin-in-schools

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2013). Longitudinal Pathways From Math Intrinsic Motivation and Achievement to Math Course Accomplishments and Educational Attainment. *Journal of Research on Educational Effectiveness*, *6*(1), 68–92. https://doi.org/10.1080/19345747.2012.698376

Haring, N. G., & Eaton, M. D. (1978). Systematic instructional procedures: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen, *The fourth R: Research in the classroom.* Charles E. Merrill.

Hasselbring, T. S., Goin, L. I., & Bransford, J. D. (1987). Developing Automaticity. *TEACHING Exceptional Children*, *19*(3), 30–32. https://doi.org/10.1177/004005998701900309

Hawkins, R. O., Collins, T., Hernan, C., & Flowers, E. (2017). Using Computer-Assisted Instruction to Build Math Fact Fluency: An Implementation Guide. *Intervention in School and Clinic*, *52*(3), 141–147. https://doi.org/10.1177/1053451216644827

Higgins, J., Chandler, J., Cumpston, J., Li, T., Page, M., & Welch, V. (Eds.). (2023). Chapter 2: Determining the scope of the review and the questions it will address. In *Cochrane*

Handbook for Systematic Reviews of Interventions (6.4). Cochrane. https://training.cochrane.org/handbook/current/chapter-02

Hume, S., Brown, S. R., & Mahtani, K. R. (2023). School closures during COVID-19: An overview of systematic reviews. *BMJ Evidence-Based Medicine*, *28*(3), 164–174. https://doi.org/10.1136/bmjebm-2022-112085

Jay, T., Habgood, J., Mees, M., & Howard-Jones, P. (2019). Game-Based Training to Promote Arithmetic Fluency. *FRONTIERS IN EDUCATION*, *4*, 4–118. https://doi.org/10.3389/feduc.2019.00118

Korakakis, V., Whiteley, R., Tzavara, A., & Malliaropoulos, N. (2018). The effectiveness of extracorporeal shockwave therapy in common lower limb conditions: A systematic review including quantification of patient-rated pain reduction. *British Journal of Sports Medicine*, *52*(6), 387–407. https://doi.org/10.1136/bjsports-2016-097347

Kromminga, K. R., & Codding, R. S. (2023). The impact of intervention modality on students' multiplication fact fluency. *Psychology in the Schools*, 1–23. https://doi.org/10.1002/pits.23054

Maths Circle Ltd. (2024). *Times Tables Rock Stars – Times Tables Rock Stars*. https://ttrockstars.com/

Moyer-Packenham, P. S., Lommatsch, C. W., Litster, K., Ashby, J., Bullock, E. K., Roxburgh, A. L., Shumway, J. F., Speed, E., Covington, B., Hartmann, C., Clarke-Midura, J., Skaria, J., Westenskow, A., MacDonald, B., Symanzik, J., & Jordan, K. (2019). How design features in digital math games support learning and mathematics connections. *Computers in Human Behavior*, *91*, 316–332. https://doi.org/10.1016/j.chb.2018.09.036

NCETM. (2017). *Five Big Ideas in Teaching for Mastery*. National Centre for Excellence in the Teaching of Mathematics (NCETM). https://www.ncetm.org.uk/teaching-for-mastery/mastery-explained/five-big-ideas-in-teaching-for-mastery/

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, *10*(3), 233–240. https://doi.org/10.1111/jebm.12266

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, *372*, n71. https://doi.org/10.1136/bmj.n71

Pegram, J., Watkins, R. C., Hoerger, M., & Hughes, J. C. (2022). Assessing the range and evidence-base of interventions in a cluster of schools. *Review of Education*, *10*(1), e3336. https://doi.org/10.1002/rev3.3336

Reddy, B. (2010). *Times Tables Rock Stars*. Maths Circle Ltd.

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology*, *60*, 101832. https://doi.org/10.1016/j.cedpsych.2019.101832

Skaalvik, E. M., Federici, R. A., & Klassen, R. M. (2015). Mathematics achievement and self-efficacy: Relations with motivation for mathematics. *International Journal of Educational Research*, *72*, 129–136. https://doi.org/10.1016/j.ijer.2015.06.008

Smith, C. L. (2010). *Examining the effectiveness of peer-tutoring and computer-aided instruction for mastery of multiplication facts* [Ed.D., Northcentral University]. https://www.proquest.com/docview/275860889/abstract/782DFA1E6A3C4753PQ/1

Standards and Testing Agency. (2022). *Multiplication tables check assessment framework: From academic year 2021/2022 For Test Developers*. UK Government. https://assets.publishing.service.gov.uk/media/62384958e90e0779a5e700c3/2018_MTC_assessment_framework_PDFA_updated_for_2022.pdf

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Ward, H. (2015). Turning lessons up to 11 with rock star times tables. *The Times Educational Supplement*, *5143*, 14–15.

Woodward, J. (2006). Developing Automaticity in Multiplication Facts: Integrating Strategy Instruction with Timed Practice Drills. *Learning Disability Quarterly*, *29*(4), 269–289. https://doi.org/10.2307/30035554

Appendix A

Systematic Search Strategy

After an initial scoping search using Google Scholar, Web of Science, Educational Resources Information Center (ERIC), and APA Psych Info returned no results for the term "Times Tables Rock Stars", additional scoping searches were conducted for the terms 'gamification', 'digital app', 'multiplication fluency', 'times tables', and 'digital intervention'. Due to there being no existing studies looking into the impact of Times Tables Rockstars the PICO (Higgins et al., 2023) framework was adopted to create a specific question for the systematic literature review, as recommended by the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2023). Table A1 shows the key terms determined through the PICO framework which resulted in the following specific question: 'Are digital games effective in increasing multiplication fact fluency in school aged children and young people?'

Table A1

PICO (Higgins et al., 2023) Systematic Review Question Formulation and Related Search Terms

PICO Term	What This Means for TTRS	Search Terms with Boolean Operators
Population	School aged pupils, primary or secondary school aged, over 6 years old	(pupil OR student OR children OR adolescents OR youth OR child OR teenager)AND
Intervention	Game for practicing times tables using technology/website/app	(gamification OR gaming OR digital OR app OR online OR "computer assisted" OR "computer aided")AND
Comparison(s)	N/A or could be traditional teaching methods	
Outcome	Automatic recall of multiplication facts. Increased multiplication fluency.	("times tables" OR multiplication)AND(fluency OR fluent OR recall OR acquisition)

A search was then conducted using search terms identified from synonyms of key terms in the specific question (Table A1) using Boolean operators OR and AND. The databases APA PsychInfo, ERIC, and Web of Science, were identified as most relevant for education interventions, and therefore used in the search with results restricted to the time period 1st January 2008 to 1st September 2023 to match the inclusion and exclusion criteria (Table A2). An additional search to look for grey literature in order to limit publication bias and widen the evidence base for the systematic review (Paez, 2017) was conducted using the same time period restrictions and the addition of abstract only criteria using the ProQuest Dissertation and Theses Database. A hand search of included papers reference lists was also conducted to ensure no additional relevant studies were missed, however no additional papers were included from this search.

Table A2

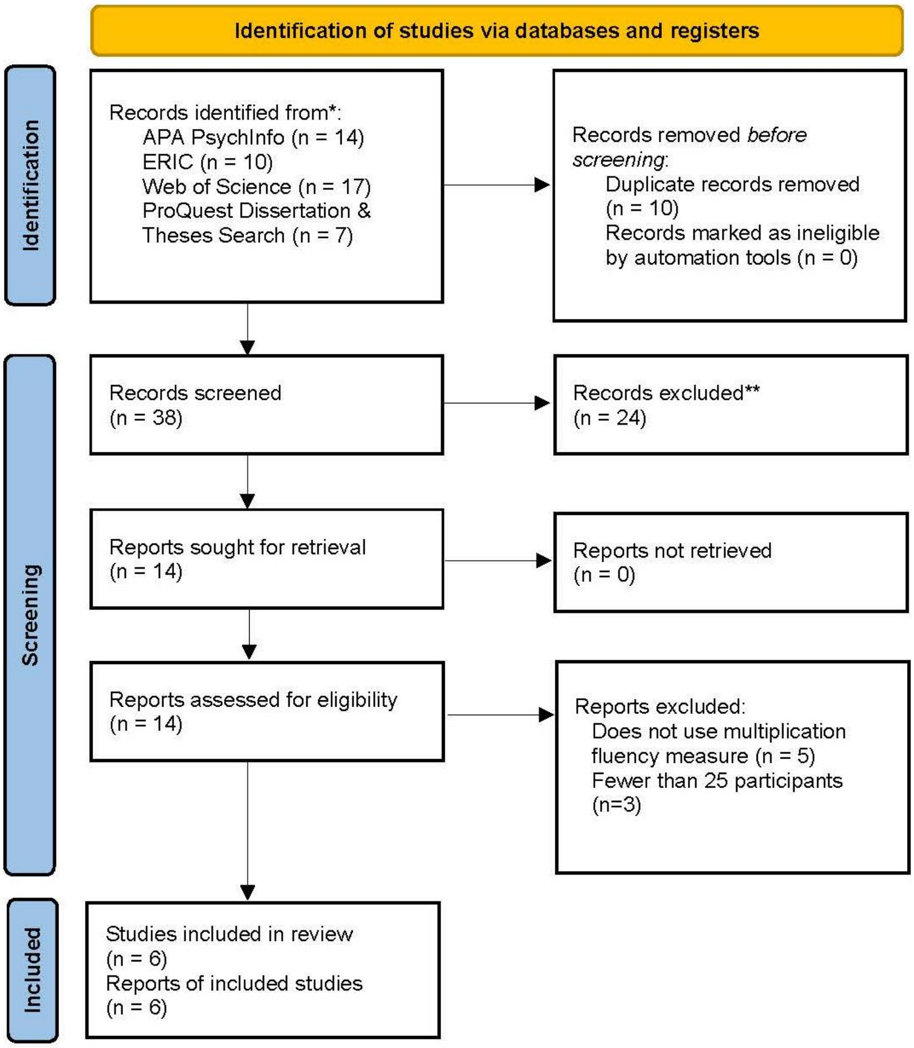
Systematic Literature Review Inclusion Criteria

Inclusion Criteria	Exclusion Criteria
Study written in English	Study written in language other than English
Participants must be school aged and at least six years old	Participants under six or over 18 years of age
Study must be published within the past 15 years	Study older than 15 years
Must have a measure that includes assessment of multiplication fact fluency	Does not include multiplication fact fluency measure
Must assess a digital intervention focused on building multiplication fact fluency	Does not assess a digital multiplication fluency building intervention
Must be an empirical paper or study	Not an empirical paper or study, e.g. conference proceedings or review article
Must have some element of quantitative analysis	No quantitative analysis
Must have more than 25 participants	Fewer than 25 participants

A total of 48 studies were found through the initial database searches, of which 10 were duplicates and therefore removed before screening. Figure A1 shows the PRISMA (Page et al., 2021) flow diagram indicating the process by which papers were screened and eventually included in the systematic review. Initial screening involved looking through titles and abstracts to identify all papers that met the exclusion criteria.

Figure A1

PRISMA (Page et al., 2021) Flow Diagram



*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

After the initial screening removed 24 papers, the remaining 14 were successfully retrieved and the full paper examined in line with the inclusion criteria. This process resulted in eight additional papers being removed, and the remaining six were then quality assessed using an adapted version of the Downs and Black (1998) checklist (Korakakis et al., 2018), a summary of which can be found in Appendix B. This quality assurance checklist was chosen as it focuses on quantitative studies with the adapted version providing space for comments and an amended scoring system to allow a quick overview comparison of study quality. Finally, a data extraction table was created (Appendix C) to support the synthesis of the evidence in answer to the specific systematic review question.

Appendix B

Quality Appraisal of Included Studies using Downs and Black (1998) Checklist Adapted by Korakakis et al., (2018)

"In the present version of the checklist we modified the scoring of item 27 that refers to the power of the study. Instead of rating according to an available range of study powers, we rated whether the study or not performed power calculation. Accordingly the maximum score for item 27 was 1 (a power analysis was conducted) instead of 5 and thus the highest possible score for the checklist was 28 (instead of 32). Downs and Black score ranges were given corresponding quality levels as previously reported (Hooper, Jutai, Strong, & Russell-Minda, 2008): excellent (26-28); good (20-25); fair (15-19); and poor (≤14)." (Korakakis et al., 2018)

	Agee (2019)	**Berrett and Carter (2018)**	**Denham (2013)**	**Jay et al. (2019)**	**Kromminga and Codding (2023)**	**Smith (2010)**
1. Is the hypothesis/aim/objective of the study clearly described?	Yes	Yes	Yes	Yes	Yes	Yes
2. Are the main outcomes to be measured clearly described in the Introduction or Methods section?	Yes	Yes	Yes	Yes	Yes	Yes
3. Are the characteristics of the patients included in the study clearly described?	Yes	Yes	Unknown	Unknown	Yes	Yes
4. Are the interventions of interest clearly described?	No	Yes	Yes	Yes	Yes	Yes
5. Are the distributions of principal confounders in each group of subjects to be compared clearly described?	No	Yes	No	Unknown	Yes	Yes
6. Are the main findings of the study clearly described?	Yes	Yes	Yes	Yes	Yes	Yes
7. Does the study provide estimates of the random variability in the data for the main outcomes?	Yes	Yes	Yes	Yes	Yes	Yes
8. Have all important adverse events that may be a consequence of the intervention been reported?	No	Yes	No	No	Yes	Unknown
9. Have the characteristics of patients lost to follow-up been described?	No	Yes	No	No	Yes	Unknown
10. Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	Yes	None reported	No	Yes	Yes	No
11. Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	Yes	Unknown	Unknown	Unknown	No	Yes

	Agee (2019)	Berrett and Carter (2018)	Denham (2013)	Jay et al. (2019)	Kromminga and Codding (2023)	Smith (2010)
12. Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	Unknown	Unknown	Unknown	Unknown	Unknown	Yes
13. Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?	Yes	Yes	Yes	Yes	Yes	Yes
14. Was an attempt made to blind study subjects to the intervention they have received?	No	No	No	No	No	No
15. Was an attempt made to blind those measuring the main outcomes of the intervention?	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
16. If any of the results of the study were based on "data dredging", was this made clear?	Yes	Yes	Unknown	Yes	Yes	Yes
17. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?	Yes	Yes	Yes	Yes	Yes	Yes
18. Were the statistical tests used to assess the main outcomes appropriate?	Yes	Yes	Unknown	Yes	Yes	Yes
19. Was compliance with the intervention/s reliable?	Yes	Yes	Unknown	Yes	Yes	Yes
20. Were the main outcome measures used accurate (valid and reliable)?	Yes	Yes	Unknown	Yes	Yes	Yes
21. Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	Yes	Yes	Yes	Yes	Yes	Yes
22. Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	Yes	Yes	Yes	Unknown	Yes	Yes
23. Were study subjects randomised to intervention groups?	No	Yes	Yes	Unknown	Yes	Yes
24. Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	No	Unknown	Unknown	No	Unknown	No

	Agee (2019)	Berrett and Carter (2018)	Denham (2013)	Jay et al. (2019)	Kromminga and Codding (2023)	Smith (2010)
25. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	No	Yes	Unknown	No	No	No
26. Were losses of patients to follow-up taken into account?	No	Yes	Unknown	No	No	Unknown
27. Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%? Sample sizes have been calculated to detect a difference of x% and y%.	No	Unknown	No	Unknown	No	No
Total Score based on Korakakis et al., (2018)	15	21	10	13	20	19
Overall Rating based on Korakakis et al., (2018)	Fair	Good	Poor	Poor	Good	Fair

Appendix C

Data Extraction Table

Study	Design	Participants	Intervention	Comparison/Control	Measures	Results	Limitations
Agee (2019) – Doctoral Thesis	Quantitative quasi-experimental	114 children aged 12-14 from one 'traditional fairly rural' middle school in USA. Participants described as 91% Caucasian and approximately 17% with a recognised disability.	The online fact fluency platform of Moby Max a web based interactive learning platform. Practicing multiplication facts from 0-12 multiplication tables. 2 intervention groups, one used Moby Max for 10 days, the other for 20 days. During intervention time students spent 20 minutes every 'regular school day'.	Two digital intervention groups compared with no treatment control. Measures from all groups taken pre-test, mid-test (after the 10 day group completed intervention) and post-test (after the 20 day group completed intervention.	Researcher generated paper and pencil multiplication fact fluency assessment using online random worksheet generator. Score = number of correct answers in 1min Adapted Multiplication Attitude Survey to measure mathematical self-efficacy.	Mid test results showed that the 10 day intervention group had significantly higher scores than the control group, where at baseline there had been no significant difference between groups. However, the 10 day group also had significantly higher scores than the 20 day group at mid-test and at this point both groups had had exactly the same level of intervention. At post-test there was no significant difference between groups – however due to there being a significant difference post-test this could mean that there was a significant difference in level of improvement for the extended group. No significant difference between rate of improvement for 'general education' and 'special education' students according to ANOVA – but descriptive statistics showed 'general education' group made more gains than 'special education' group, although standard deviation was very high for 'special education group'. There was a significant increase in pupil's mathematics self-efficacy	Any pupil with missing data was excluded from the final analysis. There was a statistically significant difference between the control group and the intervention extended group at pre-test with the control group scoring significantly higher on multiplication fluency. This may mean they were not a representative sample and did not provide an accurate control comparison. Intervention took place during maths lesson time rather than in supplement to it. Looked at scores at mid- and post test rather than rate of improvement in analysis between groups. ANOVA of participant characteristic comparisons underpowered due to low numbers in certain groups. Unclear what the control group focused on in maths lessons during the intervention period.

Study	Design	Participants	Intervention	Comparison/Control	Measures	Results	Limitations
						from pre to post test after using the digital intervention.	
Berrett and Carter (2018)	Multiple baseline across groups	63 children aged 8-9 from a charter school with 'typical' levels of attainment for the local area in 'suburban western USA'. Half participants 'near' or 'below' expected maths attainment levels (n=32) the rest at or above expected levels (n=31).	Timez Attack, part of the wider Imagine Math Facts program. 3D world the pupil navigates through, answering multiplication facts to progress through the game. Staggard baseline design meaning one group had intervention for 7 ½ weeks, one had intervention for 6 ½ weeks, and the third had intervention for 5 ½ weeks. During intervention phase pupils played the game for 20-30 minutes twice a week in school, they were instructed not to play the game outside of this time.	Staggard baseline design, meaning all groups had baseline period, intervention period, and post-test period. During the 'baseline' period participants played on an English language and literacy computer game.	Paper and pencil 30 question assessment of multiplication fact fluency created by the researchers using an online worksheet generator – 22 different worksheets completed throughout the course of the research, 2 per week completed immediately before playing game. All comprised of single digit (i.e. 1-9) multiplication table facts, with pupils being given 1 minute to answer as many as possible. Score = number of correct answers in 1 min.	Missing data was delt with by mean imputation: adding the mean average score during intervention and baseline. 62% of students had some missing data, with the maximum number of data points missed being 5. (potentially due to high level of data collection?) Visual inspection of graphs showed an interaction effect at transition from baseline to intervention phase. However, consistency in baseline was not as expected, this was explained as resulting from assessment 6 and 8 containing a higher number of 2x and 1x multiplication fact questions, meaning these were easier for all participants to gain high scores in whether they were in the intervention or baseline control phase. A post-hoc analysis which was not an original study objective, looked at the relationship between prior ability level and gains made through intervention. Found a correlation between prior ability level and gains made, with those with higher prior ability level making most gains. Although all groups made gains throughout the intervention, the group with 'below proficient' levels of mathematical ability before the intervention	All participants from one school, which was a charter school so may not represent wider population in area or internationally. Does not specify if participant group matched school demographics as described. Assessments were not checked for uniformity before the research began meaning some inconsistency in potential level of difficulty occurred, impacting results. Measures limited pupils to 30 questions in 1 min and for some pupils this led to a ceiling effect, meaning the true impact on their improvement was not recorded.

Study	Design	Participants	Intervention	Comparison/Control	Measures	Results	Limitations
						increased by 5 points, where as the group with above proficient levels at baseline increased by 11.1 points. Non-overlap of all pairs (NAP) effect size calculation showed a strong treatment effect for all groups, with the group who spent the longest time in the intervention phase having the highest effect size.	
Denham (2013)	Pre/post intervention mode comparison	29 children aged 8-10 from one 'predominantly African-American' elementary school in Southeast USA. No learning disabled or gifted children included.	Lerpz: Escape from Goldac game. Designed specifically for the research. Pupils must solve multiplication facts in a set amount of time to unlock 'force fields'. One version of the game had 'heads-up-display' which gave pupils strategies for solving problems, the other did not. Intervention lasted for 2 days, 30 minutes game play during each day, in a 'computer lab' (unclear if in school).	Compared two versions of the same digital game intervention, one with timed drills only and one with additional strategies available to pupils.	Timed paper based single-digit multiplication fact test: 100 questions, score = how long it took to answer all 100 questions (unclear if included way of recording incorrect answers) Conceptual understanding of multiplication test: word problems based on existing research. 21 questions and pupils also asked to describe multiplication in their own words.	Combined scores for all participants showed significant decrease in time taken to answer 100 multiplication questions. Between groups comparison found no significant difference. Post-hoc analysis for gender difference found no significant difference. No significant difference was found between groups for the post-test conceptual understanding of multiplication tests measure. Participants reported enjoying the game and feeling that it was beneficial.	Pre-test scores indicate quite average level of fluency already with quite a wide range judging by standard deviation. Unclear if tests measured speed and accuracy or just speed. Small sample size means ANOVA is likely to have been under powered and would not necessarily have picked up medium or small effects. Intervention only ran for two days with total of 1 hour of game play, which may not be long enough for the impact of the intervention to be seen.

Study	Design	Participants	Intervention	Comparison/Control	Measures	Results	Limitations
					Game play survey – a questionnaire on pupils experiences/opinions of the game.		
Jay et al. (2019)	Intervention control (pre/post test) crossover design	90 children aged 7-8, from 4 classes in 2 UK schools with overall 'average' socio-economic status and maths attainment. Children not eligible if they had learning difficulties.	RAIDING Game where pupils use touch-screen controls to move around a 'outer space' environment earning credits to move through game by completing multiplication or number-bond questions. Pupils can choose which activities to take part in. Pupils played the game for 20 minutes per day for two weeks during the intervention period.	Control group with lessons as usual	Westwood 1-min basic facts test: complete as many single digit addition problems in 1 min, followed by subtraction, multiplication, and division: score = number of correct answers in 1 min Data also collected from game play	At mid-test both control and intervention group improved their scores on the arithmetic fluency tests, but the intervention group had significantly greater improvements than the control group, with an effect size indicating a high impact. From mid to post test (when the original control group became the intervention group and the intervention group became the control) there was also a significant difference between groups with the intervention group having significantly higher score improvement than the control group. However, the 'control' group did maintain their mid-test scores, indicating a potential lasting impact from the digital intervention. More 'effort' within the game play corelated with higher improvements in progress in arithmetic fluency. Playing more times tables tasks improved times tables and division fluency over playing 'number-bonds' tasks. Engagement with the game was seen to be approximately the same	Excluded participants who missed sessions from final analysis and did not explore potential reasons why these children were unable to take part in all sessions (e.g. were these the children with highest level of need for intervention?) Pupil's able to choose whether to engage in multiplication tables or number-bonds, therefore correlation with increased time on multiplication over number-bonds could be due to pupils feeling less confident in multiplication avoiding these questions. Unclear what 'lessons as usual' were.

Study	Design	Participants	Intervention	Comparison/Control	Measures	Results	Limitations
						for those with lower and higher pre-test scores. However, higher pre-test scores correlated with higher number of 'rocks selected' and higher number of correct answers in game, indicating a relationship between performance in the game and pre-test performance.	
Kromminga and Codding (2023)	Randomised digital, traditional, and control comparison pre/post test	72 children aged 8-9 from 5 classes across 2 public schools in suburban school districts in midwestern USA. Participants described as 51.2% male, 81.94% white, 80.8% no eligible for reduced meals, 93.2% native English speakers, 87.8% not eligible for special education support.	Mental Math Cards Challenge app, drill and practice digital flash cards. Participants used the app to practice multiplication tables from 0-12. 20-minute sessions over 4 days a week for 5 weeks, in regular maths lesson time, with all study groups being present in the same classroom during the intervention time. 10 minutes of the intervention session was spent playing the game.	Two comparison groups.: 1. Reciprocal Peer Tutoring (RPT): pupils matched in pairs by scores on initial pre-test with class split in half into highest scorers and lowest scorers and then matched in sequence: highest of high scorers with highest of low scorers, and so on. During intervention pairs took turns being the tutor and tutee for 5 min each, showing flash cards to each other and providing corrective feedback and praise. 2. Combined digital and peer tuition. Three of four sessions weekly on	2 min paper and pencil curriculum based measure (CBM) 2 min iPad delivered CMB Both CMB = x3 pre-test and x3 post-test, with median score used in analysis. Each CMB contained 84 multiplication facts to answer in 2 minutes. Mathematics fluency and mathematics calculation subtests of the Woodcock-Johnson 4th ed. Test of achievement Engagement was monitored	There was a strong correlation between the ipad delivered and paper and pencil CMB indicating they were very similar. There was a significant main effect for treatment group in the iPad CBM post-test, with further analysis indicating the significant difference was between the iPad group and the combined group only (in favour of the iPad group). There were no other significant differences found between groups on any measure according to ANCOVAs. Descriptive statistics showed all groups made improvements in all measures from pre-test to post-test. Although no significance analysis took place, the iPad group made just over 2 points more progress on the ipad delivered CBM than the other two groups, where as on the pencil and paper CBM all groups improvement scores were within 1.15 points of each other. Meaning there may have been some benefit to the iPad	Group size means ANCOVA under powered, so could have missed significant results with small and medium effect size. Students who missed sessions due to being taken out to 'receive special education services' were excluded from the analysis. Measures of engagement and off-task behaviour through momentary sampling may have not reflected actual level of engagement with activity as pupils looking at each other and cards may have appeared to be 'on-task' when they could have been working more slowly or talked about other things not apparent to the researchers from the observation. It may also be that iPad group needed to look away from the screen more due to eye strain but

Study	Design	Participants	Intervention	Comparison/Control	Measures	Results	Limitations
				iPad, fourth session peer tuition.	through momentary sampling observations. Adapted Kid's Intervention Profile to measure pupils views on the intervention post-test.	group in terms of practicing the mode of assessment. However, small sample size, and the fact that the combined group who also used iPads 3 days a week did not have same gains means it may not generalise. As a whole participants scores were significantly higher on the paper and pencil CMB than the ipad CMB at both pre and post test. Momentary sampling showed that students in the peer condition had significantly less off-task behaviour and more engagement time than those in the iPad condition. The iPad condition had the highest average rate of acceptability as rated by pupils, however there was no significant difference between scores and all groups reported high levels of acceptability.	when they were 'on-task' they were working at a more intense rate than the peer tutoring group, as there were not significant differences between the groups in terms of improvement in scores.
Smith (2010) – Doctoral Thesis	Pre/post test, control group experimental design	116 children aged 9-11 who 'lacked automaticity with one digit by one-digit multiplication facts' (as measured by scores of under 80 correct answers during 2 minute 100 question multiplication facts test) from 10 classes across 6 elementary	FASTT math software. Programme begins with a test of students current level of knowledge presenting multiplication facts with 1.25 seconds given for answers (after first assessing typing speed), then during game play a combination of known and	Mastering Math Facts Peer Tutoring Programme group and a control 'teaching as usual' group. Due to the way the research was designed, participants were allocated to a group before parental consent was returned, resulting in the control group having significantly	2 minute paper and pencil single-digit multiplication facts test. Score = number of correct answers in 2 minutes.	All groups made gains from pre- to post- test. However FASTT Math resulted in significant gains compared to control group, meaning it is an effective method for supporting pupils in need of additional help to improve fluency in their times tables. Computer aided group had more significant gains than peer tutoring so there is potential for this to be a more useful intervention for those who are not at expected levels than a peer support intervention – however the difference could also	Small sample size for SEN pupils means ANCOVA under powered

Study	Design	Participants	Intervention	Comparison/Control	Measures	Results	Limitations
		schools in southwest Ohio, USA. Participants described as 93.1% white, 17.2% disabled, 40.5% entitled to reduced meals.	unknown multiplication facts are given. Intervention completed for 10 minutes per day over 6 weeks with participants from all groups present in regular classrooms during intervention time.	fewer participants (n=28) than the digital (n=42) or peer (n=46) intervention groups.		be due to intervention fidelity/human error in this study.	