Precision Teaching for maths: An academic critique

Nicole Harris[1] & Sarah Wright[1]

[1]School of Psychology, University of Southampton, Southampton, SO17 1BJ

*Corresponding author:

Dr Nicole Harris, Trainee Educational Psychologist, University of Southampton, UK.

Email: n.m.harris@soton.ac.uk

Precision Teaching for maths: An academic critique

Children are not retarded. Only their behavior in average environments is sometimes retarded. In fact, it is modern science's ability to design suitable environments for these children that is retarded…The purpose of this paper is to suggest techniques of designing prosthetic environments for maximizing the behavioral efficiency of exceptional children who show deficits when forced to behave in average environments.

Classrooms should be tailored to children- not children adjusted to classrooms…In prosthetic environments tailored to their skills, exceptional children will behave adjustively, efficiently, and with full human dignity.

(Lindsley, 1964, p. 62, 80)

It is with these words, perhaps dated now in their terminology and acceptability, that Ogden Lindsley (1964) introduced and concluded his seminal paper on the methodology that would become known as Precision Teaching (PT). In this academic critique, I will explain what PT is, who it is for and explore its theoretical underpinnings. I will then report the findings from a systematic search of the literature that aimed to answer the question 'Is Precision Teaching effective at improving maths skills in school aged children and young people'.

What is Precision Teaching?

PT can be explained as a way of trying to establish 'what teaches best', rather than it being a teaching method in and of itself (Raybould & Solity, 1982). It relies on giving daily feedback, which helps establish the effectiveness of teacher instruction. PT provides the techniques that enable the teacher to measure, chart and evaluate progress (Raybould & Solity, 1982), but leaves the 'how' of the teaching to the teacher. Crucially, if a child is not making progress, the expectation is for the teacher to examine and alter their way of teaching, rather than assuming a difficulty in the child's ability to learn (Raybould & Solity, 1982). One of Lindsley's (1995) founding principles of PT is that 'the child knows best' and as such will, through working with the teacher, help co-construct the most effective teaching-learning strategies.

There are typically 5 agreed steps to PT (Evans et al., 2021; Johnson & Street, 2012):

- teacher identifies a learning objective (the pinpoint);
- teacher arranges materials and procedures to learn and practice the pinpoint;
- teachers calculates performance frequency and charts performance on standard celeration chart (SCC);
- teacher and learner review performance trends on the SCC and decide how to improve performance; and
- try, try again.

Pinpoints typically set out the mode in which the learner interacts with the material and the behavioural response, including a targeted performance standard, for instance hear word/say definition (the learning channel), 50-60 per minute (performance standard) (Gist & Bulla, 2022).

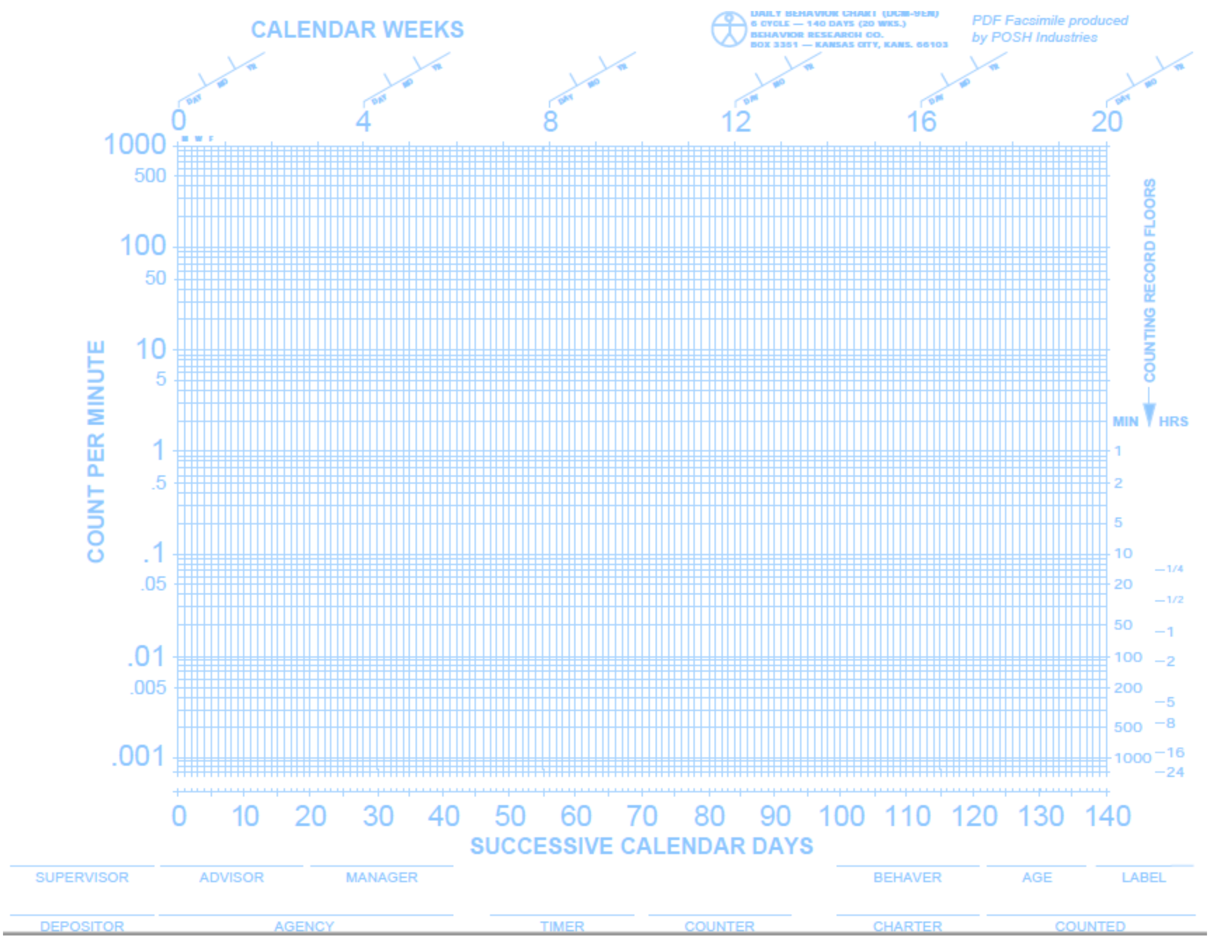
Whilst some have claimed that PT is a misnomer and it is better described as precision recording, Raybould & Solity (1988) defended the name, explaining that it embodies an approach and attitude to teaching basic skills which includes identifying precise teaching targets, breaking them down into subskills and changing the teaching programme as required. It is therefore not simply a process of measurement and recording. Evans et al. (2021), recognising the long history and subsequent growth and evolution of PT set out to create an updated definition, which included all its critical features (p. 561):

PT is a system for precisely defining and continuously measuring dimensional features of behaviour and analyzing behavioral data on the SCC to make timely and effective data-based decisions to accelerate behavioral repertoires.

They specifically chose the term system, as it refers to a set of specific elements which work together to achieve a set purpose (Evans et al., 2021).

Figure 1

A sample standard celeration chart (SCC) (University of Washington, n.d.)



Frequency measurements and the SCC are key concepts within PT. Charting frequency has been shown to be ten to a hundred times more sensitive to changes in performance than percentage-based measures (Lindsley, 1962, cited in Lindsley, 1991).

Moreover, it has been shown to be linked to the probability of future performance (Binder, 1996; Lindsley, 1992, cited in Evans et al., 2021). A frequency measure determines how many times the pupil has been able to carry out the desired behaviour in a set time, for example correct response per minute. The SCC is a logarithmic chart (see Figure 1) that presents a visual representation of progress. *Celeration* is a word derived from the root words of acceleration and deceleration, and is the term used to refer to the change in learning (Calkin, 2011). Plotting celeration on the SCC enables both the learner and teacher to see when progress has been accelerated, through a change in slope of the line charted (Lindsley, 1991) or if it has not, and an alternative teaching method should be tried. The SCC is designed such that a line drawn from the lower left to the upper right represents a doubling of frequency over one week. The slope as measured from corner to corner is therefore referred to as a celeration of times two per week (Lindsley, 1991). The slope measures three dimensions- number per minute per week- and is a measure of learning rate (Binder, 1996; Lindsley, 1991).

Who is Precision Teaching For?

In their 1982 paper, Raybould and Solity suggested that children with learning difficulties benefited from the structure of PT. They expanded this to all children who have difficulties with basic skills six years later (Raybould & Solity, 1988); they did not advocate its use for all children. PT, however, subsequently moved into the mainstream classroom as whole-class provision (Binder, 1996), and Evans et al. (2021) show that PT as a system lends itself to accelerating behavioural repertoires in a diversity of individuals and arenas, giving examples of application by football players, piano teachers and individuals with depression. PT has moved beyond Lindsley's classroom to have a much wider application and is used in a variety of settings to teach a myriad of populations a wide range of behaviours (Ramey et al., 2016).

Theoretical Underpinnings

PT originated in Skinner’s laboratory when Lindsley sought to extend their lab-based work on animals, to people (Evans et al., 2021). As such, PT is rooted in behavioural principles, specifically free operant conditioning. PT assumes that behaviour responds predictably to changes in environmental variables, that behaviours are observable and can therefore be reliably and objectively measured, that frequency is a dimension of behaviour (Lindsley, 1991), and that performance frequency can be increased (McTiernan et al., 2022).

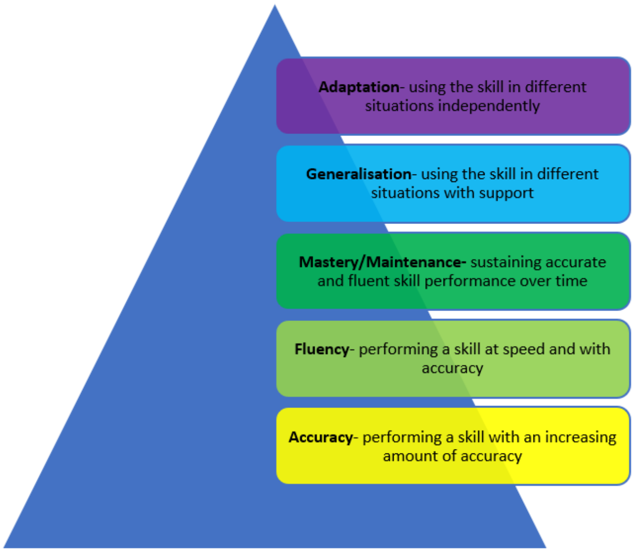
Lindsley (1991) believed that frequency was more than a measure of behaviour, that it was a dimension of behaviour and that a behaviour was not accurately described until its frequency was stated. By extension, in changing the frequency, the behaviour changed. The use of frequency in PT is therefore much more than a simple metric- it is core to behavioural change.

PT sits within the theory of behavioural fluency, which emerged from free-operant conditioning (Binder, 1996). Binder (1988) defines behavioural fluency as a fluid combination of accuracy and speed, indicative of a competent performance (Binder, 1988). It was the work of precision teachers differentiating the stages of the learning process that supported the development Haring and Eaton’s (1978) Learning Hierarchy (Binder, 1996) (see Figure 2), which establishes accuracy and fluency as the base skills on which all further learning rests. As stated by (Gallagher et al., 2006), there is evidence to suggest that fluency training assists higher academic achievement for all learners, however, key research questions remain unanswered, for instance around the endurance of the learned material and limitations of the PT technique (Gallagher et al., 2006; Raybould & Solity, 1988).

The PT intervention is therefore solidly grounded in a clear and established theoretical framework. It is this very framework, however, that may have resulted in a hesitancy towards adopting PT, as many are against a behaviourist approach in teaching, wanting learning to be more discovery or experiential (le Grice et al., 1999). Despite its behaviourist underpinnings however, PT does not assume that difficulties with learning are within-child (Raybould & Solity, 1982), but asks the teacher to look to their methods, and the environment- creating the prosthetic environment of the opening quotes- to address learning.

Figure 2

The learning hierarchy (Haring & Eaton, 1978)



Impact and Effectiveness of PT

Given PTs long history, it would be easy to assume a substantial evidence base: This is not the case. This is, in no small part, due to Lindsley discouraging early precision teachers from publishing (Binder, 1996; Lindsley, 1995), as he felt that publications did not change professional behaviour enough to warrant the effort required to publish in academic journals (Lindsley, 1974, cited in Binder, 1996). This was rather perceptive, as it predicted the findings of Pegram et al., (2022) nearly 50 years later, whose research suggested that schools typically do not refer to research evidence when deciding on interventions, but rather rely on the opinions and experience of practitioners.

In addition to a lack of published data on the impact of PT, there was also a lack of controlled studies in those published (Binder, 1996). Furthermore, as PT is not a singular intervention, several different educational strategies can be used within a PT framework, for instance Direct Instruction (Lindsley, 1991) and frequency building (Gist & Bulla, 2022). This makes between-studies comparisons more complicated (McTiernan et al., 2022), however it is still possible to compare an application of PT with a control group or alternate intervention (i.e. within study comparison vs. between study).

It is this combination of factors which has created the current situation where there are still calls for the establishment of an evidence base for the effectiveness of PT (Chiesa & Robertson, 2000; Fitzgerald & Garcia, 2006; le Grice et al., 1999; Sleeman et al., 2021; Stocker et al., 2019; Strømgren et al., 2014; Vostanis et al., 2021).

Systematic Literature Search

A systematic search of the literature was therefore carried out to ascertain the efficacy of PT, addressing in particular its ability to improve basic maths skills. Mathematics is a core curricular subject and has been targeted for improved pupil performance by many countries internationally (Gallagher, 2006; le Grice et al., 1999; Strømgren et al., 2014). Interventions that improve core maths skills, with a proven, established, evidence base (Ramey et al., 2016), would therefore have a very wide application. Whilst a number of systematic reviews have recently been carried out on PT, none have focussed specifically on maths for mainstream, school-aged children (Connolly et al., 2018; Gist & Bulla, 2022; McTiernan et al., 2022; Ramey et al., 2016; Stocker et al., 2019), and many have been of low methodological quality (Ramey et al., 2016).

To help address concerns around low methodological quality, a targeted search question was used ('Is Precision Teaching effective at improving maths skills in school aged children and young people') using Boolean operators, with strict inclusion criteria and screening processes, as well as the Downs and Black quality control checklist (Downs & Black, 1998) (see Appendix A). Five studies were identified that met the criteria, and from these, core information was extracted to facilitate and standardise comparison (see Appendix B).

The chosen studies were international in scope and spanned 23 years of research. Sample sizes ranged between 23 and 48, with sample group sizes varying from five to 24.

Participant age was between eight and 13. Selection criteria varied, with the majority identifying children who were ‘falling behind/failing to keep pace’ with their peers for inclusion in the intervention group. Le Grice et al. (1999) were the only ones to use an objective measure (children scoring below the 22nd percentile) for inclusion. Equally, McTiernan et al. (2018) were objective in their inclusion criteria by including all children for whom they had assent/consent, removing any potential teacher-based selection bias. Strømgren et al. (2014) identify teacher nomination as a questionable selection criterion, as they found students in their control group with similar scores to their intervention group. Ramey et al. (2016) commented that one of the biggest methodological concerns around study rigour related to failure to conduct adequate baseline measurement: It is suggested that future studies adopt more objective measures for inclusion.

In terms of the interventions themselves, all focussed on improving core maths skills, with the majority focussing on multiplication and/or division. Most were delivered daily within the scheduled maths lesson. The length of the lesson was not always detailed, but where specified, was 20 to 40 minutes. Strømgren et al. (2014) removed their children from a variety of lessons to receive the intervention, however they did not comment on what the impact of this could be on learning in those lessons. Similarly, the children in the study by McTiernan et al. (2018) study received additional intervention sessions, but there was no detail as to when this occurred. The duration of the interventions ranged from five to 12 weeks.

None of the interventions involved 1:1 work between a teacher and child. Most involved small groups where pupils worked independently from worksheets in a folder and scored and charted their results themselves. For two studies, Chiesa and Robertson (2000) and Gallagher (2006), the children’s progress was only reviewed weekly. Whilst this way of working fits with the modernised definition of PT by Evans et al. (2021) (see section ‘What is Precision Teaching’) it seems to be missing the more collaborative approach between the teacher and learner implied in the core five steps, where the best way of learning is co-constructed. Moreover, it creates a situation where a pupil can be working on a target for five days without making progress, or stagnating on a target that they have already met. This is picked up on in Strømgren et al. (2014) where they reflect that reviewing pupil results at the end of a daily session may result in being too late or lost opportunities for effective intervention. This weekly check-in appears to be a sub-optimal application of PT with respect to pupil performance, although it could be argued that it has benefits with respect to teacher time commitments.

With regards to study design, three adopted a randomised control trial (RCT). RCTs remove bias from group allocation and support the identification of discernible and measurable intervention effects (Connolly et al., 2018). Three studies analysed their results statistically, which adds rigour to results by enabling the quantification of an effect. This was carried out most effectively by McTiernan et al., (2018) who used a MANCOVA to control for any differences that may have existed between their randomly chosen groups. Gallagher (2006) deliberately abstained from statistical analysis, stating that it did not fit with a behaviourist, inductivist approach.

Results Found in Identified Studies. Each of the studies found a positive effect of the PT intervention on their measured outcomes, whether carried out against a control or comparison intervention (see Appendix B Results, and Statistically significant results). For some this was only identified through reporting one mean being greater than another. Although standard deviations were provided, data were not graphed so it was not possible to visually check for significant difference between means. McTiernan et al. (2018) found a significant difference between pre-and post-intervention results for both their intervention and control groups showing that both groups had improved with time. Crucially, they were also able to show a statistically significant difference between groups, showing those in the PT group had made statistically significant greater gains than the control. As this study was based on a representative sample of Year 4 children (i.e. not selected as they were falling behind and compared with ‘average’ classmates), adopted an RCT design and used a MANCOVA to control for between-group differences, it is felt that this study provided the strongest evidence for the potential for PT to improve maths skills. However, as this study also used frequency building and the Morningside Maths Fluency curriculum in tandem with PT, it is not possible to dissociate the specific impact of PT.

Conclusions

This academic critique used strict inclusion and screening criteria to ensure quality research was reviewed. Many studies were excluded as they did not have a comparison or control group, making it impossible to determine whether improvements were a result of PT (Connolly et al., 2018). Despite these efforts, it has been difficult to compare or combine the findings of each of these studies to be able to speak generally to the efficacy of PT to promote maths skills (McTiernan et al., 2022). This is, in part, due to the fact that PT isn't a fixed intervention, but an approach, which means that even for methodologically strong studies such as McTiernan et al. (2018), it is not possible to identify the specific contribution of PT.

A further complication in assessing the efficacy of PT arose from nearly all the studies being group interventions, with varying levels of teacher input (i.e. ranging from once per week via updated worksheets, to small groups with a 1:2 teacher to pupil ratio). Whilst for those studies it was possible to do between-groups comparisons, they do not lend themselves to between-studies comparisons.

It must therefore be concluded that based on the review of these five studies, it is not possible to say whether PT is an effective way to support maths skills. It is possible to say that in each of the interventions reviewed here, generally, the intervention that involved PT showed greater performance gains than the control or comparison group. Given the highly specific conditions of these interventions however, it is not possible to generalise the intervention effects beyond these studies.

Implications for Professional Practice

Pegram (2022) states that the last ten years in education have seen an increased focus on evidence-based practice to support the choice of interventions that schools are choosing to implement, however that many schools and teachers still make choices based on trends and personal recommendations rather than empirical evidence. If educational professionals are to move the evaluation of interventions to something beyond the anecdotal 'it works in my class' (Pegram et al., 2022), it is suggested that more formalised and structured evaluations of interventions take place. There is a need to simplify and demystify action research so that teachers become part of the question and answer: 'Does this work? Yes, it does.' There is a need to move from the evaluation of interventions being a pursuit of the academics to a responsibility of schools and teachers- thinking that if they are going to be spending time and money delivering an intervention, they want to know that it works. This could be supported and facilitated by the creators of the intervention including a pro-forma of a simple, yet rigorous, study design that would enable teachers to collect data for their class/school, which could then form part of a larger data set. Large samples of data collected under comparable conditions, with integrity of intervention delivery, could then be analysed by the creator of the intervention, or academics, to generate empirical evidence of intervention efficacy. As stated in McTiernan (2022) in their review of PT over the last 30 years, "critical outcomes seem achievable in theory; however more rigorous, empirically-based research is needed to validate them". For an intervention that has been around for over 50 years, it seems a sad indictment of evidence-based practice that this is still being said.

References

Bennett, J. (2021). *We thinkers! series: An academic critique*. In University of Southampton Educational Psychology Research Blog. Available at: http://blog.soton.ac.uk/edpsych/?s=we+thinkers&submit=Search

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst*, *19*(2), 163–197. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2733609/pdf/behavan00020-0021.pdf/?tool=EBI

Booth, A., Noyes, J., Flemming, K., et al. (2019). Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Global Health*, *4*, 1–7. https://doi.org/10.1136/bmjgh-2018-001107

Calkin, A.B. (2011). Standard Celeration Charting. In: Goldstein, S., Naglieri, J.A. (eds) *Encyclopedia of Child Behavior and Development*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-79061-9_3145

Chiesa, M., & Robertson, A. (2000). Precision teaching and fluency training: Making maths easier for pupils and teachers. *Educational Psychology in Practice*, *16*(3), 297–310. https://doi.org/10.1080/713666088

Connolly, P., Keenan, C., & Urbanska, K. (2018a). 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

Connolly, P., Keenan, C., & Urbanska, K. (2018b). 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

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 and Community Health*, *52*, 377–384.

Evans, A. L., Bulla, A. J., & Kieta, A. R. (2021). The precision teaching system: A synthesized definition, concept analysis, and process. *Behaviour Analysis in Practice*, *14*, 559–576. https://doi.org/10.1007/s40617-020-00502-2/Published

Fitzgerald, D. L., & Garcia, H. I. (2006). Precision teaching in developmental mathematics, accelerating basic skills. *Journal of Precision Teaching and Celeration*, *22*, 11–28.

Gallagher, E. (2006). Improving a mathematical key skill using precision teaching. *Irish Educational Studies*, *25*(3), 303–319. https://doi.org/10.1080/03323310600913757

Gallagher, E., Bones, R., & Lombe, J. (2006). Precision teaching and education: Is fluency the missing link between success and failure? *Irish Educational Studies*, *25*(1), 93–105. https://doi.org/10.1080/03323310600597642

Gist, C., & Bulla, A. J. (2022). A systematic review of frequency building and precision teaching with school-aged children. *Journal of Behavioral Education*, *31*, 43–68. https://doi.org/10.1007/s10864-020-09404-3

Haring, N. G., & Eaton, M. D. (1978). Systematic instructional procedures: An instructional hierarchy. In *The fourth R: Research in the classroom* (pp. 23–40). Charles E. Merrill Publishing Co.

le Grice, B., Mabin, T., & Graham, S. (1999). *I can do maths: Changing children's mathematics percentile ranking*. http://www.aare.edu.au/99papfleg99581.htm

Lindsley, O. (1995). Precision teaching: By teachers for children. *Journal of Precision Teaching*, *12*(2), 9–17.

Lindsley, O. R. (1964). Direct measurement and prosthesis of retarded behavior. *Journal of Education*, *147*, 62–81. http://precisionteaching.pbworks.com/f/lindsley1964-DMPR.pdf

Lindsley, O. R. (1991). Precision Teaching's unique legacy from B. E Skinner. *Journal of Behavioral Education*, *1*(2), 253–266. https://link.springer.com/content/pdf/10.1007/BF00957007.pdf

McTiernan, A., Holloway, J., Leonard, C., & Healy, O. (2018). Employing precision teaching, frequency-building, and the Morningside math facts curriculum to increase fluency with addition and subtraction computation: A randomised-controlled trial. *European Journal of Behavior Analysis*, *19*(1), 90–104. https://doi.org/10.1080/15021149.2018.1438338

McTiernan, A., McCoy, A., Mendonca, J., Lydon, H., & Diffley, S. (2022). The implementation of precision teaching for the improvement of academic skills: A systematic review of the literature over thirty years. *Behavioral Interventions*, *37*(2), 505–528. https://doi.org/10.1002/bin.1852

Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*, *372*, 1–9. 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). https://doi.org/10.1002/rev3.3336

Ramey, D., Lydon, S., Healy, O., et al. (2016). A systematic review of the effectiveness of precision teaching for individuals with developmental disabilities. *Review Journal of Autism and Developmental Disorders*, *3*(3), 179–195. https://doi.org/10.1007/s40489-016-0075-z

Raybould, E. C., & Solity, J. E. (1988). Precision teaching. .. and all that! *British Journal of Special Education*, *15*(1), 32–33.

Raybould, E., & Solity, J. (1982). Teaching with precision. *Special Education: Forward Trends*, *9*(2), 9–13.

Sleeman, M., Friesen, M., Tyler-Merrick, G., & Walker, L. (2021). The effects of precision teaching and self-regulated learning on early multiplication fluency. *Journal of Behavioral Education*. http://irep.ntu.ac.uk/id/eprint/36082/1/13592_Tyler_Merrick_v2.pdf

Stocker, J. D., Schwartz, R., Kubina, R. M., Kostewicz, D., & Kozloff, M. (2019). Behavioral fluency and mathematics intervention research: A review of the last 20 years. *Behavioral Interventions*, *34*(1), 102–117. https://doi.org/10.1002/bin.1649

Strømgren, B., Berg-Mortensen, C., & Tangen, L. (2014). The use of precision teaching to teach basic math facts. *European Journal of Behavior Analysis*, *15*(2), 225–240. https://doi.org/10.1080/15021149.2014.11434723

Tiernan, A. M., McCoy, A., Mendonca, J., Lydon, H., & Diffley, S. (2022). The implementation of precision teaching for the improvement of academic skills: A systematic review of the literature over thirty years. In *Behavioral Interventions* (Vol. 37, Issue 2, 505–528). John Wiley and Sons Ltd. https://doi.org/10.1002/bin.1852

University of Washington, (n.d.) Standard celeration chart (blank) pdf. Available at: https://education.uw.edu/sites/default/files/areas/edspe/white/docs/Standard_Celeration_Chart_%28blank%29.pdf

Vostanis, A., Padden, C., Chiesa, M., Rizos, K., & Langdon, P. E. (2021). A precision teaching framework for improving mathematical skills of students with intellectual and developmental disabilities. *Journal of Behavioral Education*, *30*, 513–533. https://doi.org/10.1007/s10864-020-09394-2

Appendix A

Systematic Search

An initial scoping search of Google Scholar, ProQuest and Scopus was carried out, looking for the terms ‘precision teaching’ and ‘maths’. The PICO question framework (Population, Intervention, Comparison and Outcomes) (Richardson, 1995) was chosen to support the generation of the research question and the inclusion and exclusion criteria (see Tables A1 and A2*)*. PICO was chosen over SPICE and the more complex PerSPECTiF model as it addressed the core elements in this academic critique (Booth et al., 2019). Moreover, it was used effectively in a systematic review by Connolly et al. (2018) of evidence-based practice in education. It is suggested, however, that subsequent reviews may want to adopt a more complex, ‘systems’ perspective when reviewing PT (and so use PerSPECTiF), for instance to see to what extent a teacher’s use of PT as an intervention changes the system (the classroom and/or wider school), i.e. whether a PT approach becomes embedded in their way of teaching.

Table A1

PICO question formulation framework (Richardson, 1995)

Population	Children and young people in schools
Intervention	Precision Teaching
Comparison	Ideally to a comparison intervention or control
Outcome	Improved core maths skills

The final research question was ‘Is Precision Teaching effective at improving maths skills in school aged children and young people’. This question was used to generate the database search terms shown in Table A3.

Table A2

Inclusion and exclusion criteria for the systematic search

Inclusion Criteria	Specified as Precision Teaching
	Participants aged 16 or under
	Mainstream school-based intervention
	Full text available
	Empirical study
	Presence of a comparison or control group (this could include treatment as usual and/or wait list)
	Assessed maths skills
	Peer reviewed article, or thesis or dissertation
	Published in English
	Access full text
Exclusion Criteria	Not named as Precision Teaching (i.e. another frequency training intervention)
	Participants aged 79 or over
	Not carried out in school or mainstream provision
	No comparison group
	Assessed skill other than maths
	Qualitative study
	Published in a language other than English
	Full text not available

The specific databases searched were ERIC (Educational Resources Information Center) (ProQuest) and PsychInfo (EBSCO). The inclusion of grey literature in a review helps improve validity and decrease publication bias (Connolly et al., 2018) and so the grey literature was searched using Proquest, WWC (What Works Clearinghouse) and the EEF

(Education Endowment Foundation) (Connolly et al., 2018). Searches were limited to full texts available in English. The references of recent systematic reviews were hand searched for relevant studies to include (Gist & Bulla, 2022; Ramey et al., 2016; Stocker et al., 2019; Tiernan et al., 2022b). A forward search for articles that cited the most recent systematic reviews was carried out using the 'cited by' function in Google Scholar (Gist & Bulla, 2022). PROSPERO was searched to see if there were any prospectively registered systematic reviews.

Table A3

Search terms used in database search strategy

effectiv* OR effect OR efficacious OR improv* OR impact OR evaluat* OR achieve*
AND 'Precision Teaching'
AND 'math* skills' OR 'math* ability' OR 'math* fluency' OR 'math* aptitude'
AND child* OR youth* OR 'young people' OR 'young person' OR 'student' OR adolescen* OR teen OR pupil* OR school OR classroom OR 'primary school*' OR 'secondary school*l' OR 'high school*' OR 'elementary school*' OR 'elementary education' OR 'middle school' OR 'junior school' OR 'junior high'

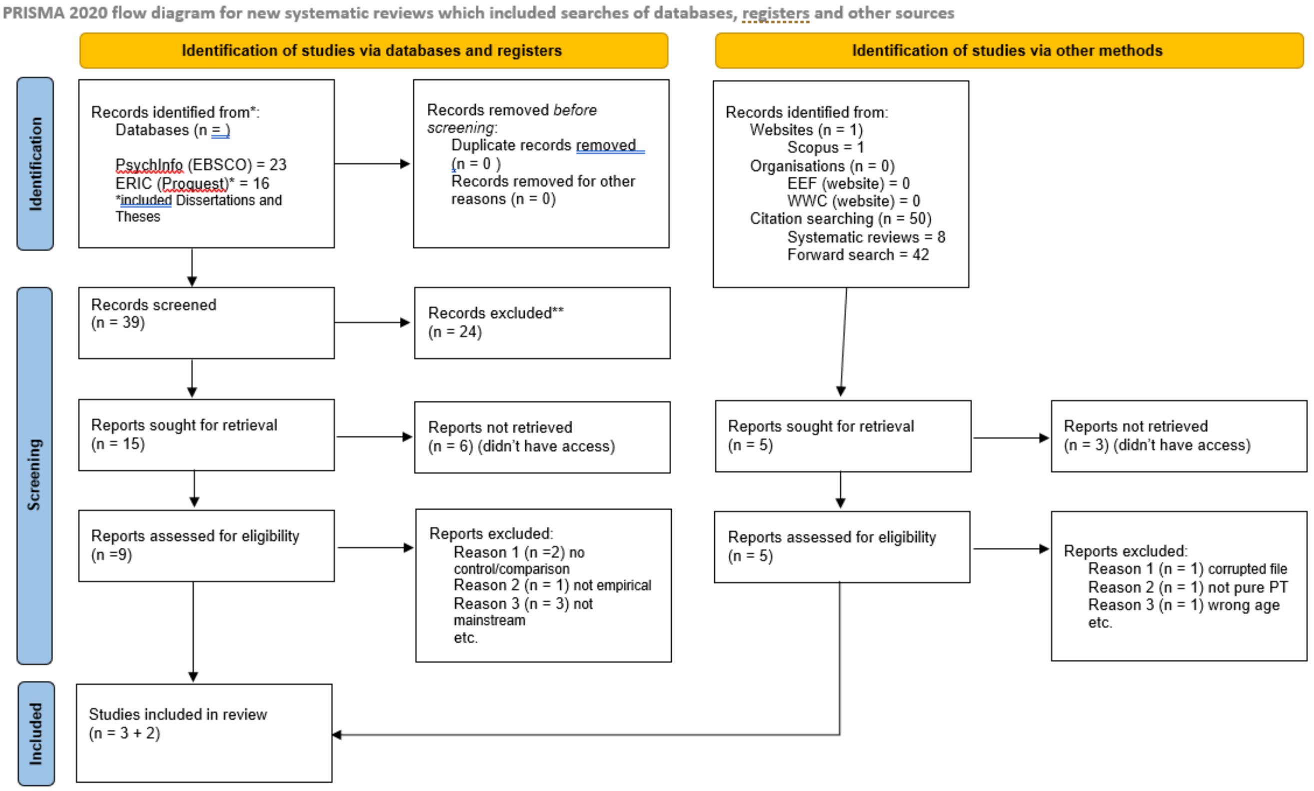
The inclusion criteria shown in Table A3 were used to filter the identified literature. This process is shown in the PRISMA flow diagram (Page et al., 2021) (Figure A1*)*. Titles and abstracts suggested by the database search were screened for relevance. Those that met the inclusion criteria then had the full paper examined to ensure inclusion criteria were fully met.

The Downs and Black (Downs & Black, 1998) checklist was used to assess the quality of the papers. Downs and Black was used as it assesses both randomised and non-randomised quantitative studies. Moreover, it can be used to generate an overall quality score, thereby facilitating the quantification of study quality.

Finally, relevant data from each study were extracted to facilitate analysis of the efficacy of the PT intervention. Included studies were summarised in relation to: geographic location; age of participants; sample size; selection criteria; intervention; frequency of intervention; duration of intervention; who delivered the intervention; training given; individual or group; group sizes; control or comparison study; measured outcomes; design; RCT; results; statistical analysis; statistical results; longitudinal results; conclusions and follow-up suggestions made by the authors (Bennett, 2021; Connolly et al., 2018; McTiernan et al., 2022) (see Appendix B).

Figure A1

PRISMA flow diagram (Page et al., 2021)



*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/

Appendix B

Data Extraction Table

Appendix B can be requested directly from the authors: n.m.harris@soton.ac.uk or s.f.wright@soton.ac.uk . Appendix B summarises key information from each of the 5 studies, namely: geographic location; age of participants; sample size; selection criteria; intervention; frequency of intervention; duration of intervention; who delivered the intervention; training given; individual or group; group sizes; control or comparison study; measured outcomes; design; RCT; results; statistical analysis; statistical results; longitudinal results; conclusions and follow-up suggestions made by the authors