

Case Study 1: An Evidence-Based Practice Review Report

Theme: School Based Interventions for Learning

How effective is Precision Teaching in improving the word reading skills of school aged students in the United Kingdom and Ireland?

Section 1: Summary

This literature review aims to evaluate the effectiveness of Precision Teaching as a school based intervention to improve the word reading skills of school aged students.

Precision Teaching is a systematic and precise method of evaluating whether a student has or has not acquired knowledge and skills within a particular curriculum area. Precision Teaching aims for students to acquire the skills of mastery, maintenance and generalisation (Binder, 1988) within a particular curricular area, highlighting the importance of students becoming fluent in a particular domain, for example word reading or multiplication.

A systematic literature search was carried out using four databases (PsychINFO, ERIC, Medline and Web of Science) alongside searching through the Journal of Precision teaching and carrying out an ancestral search. Six studies met the inclusion criteria and were reviewed using either the Horner et al. (2005) protocol (Quality Indicators: Within Single-Subject Research, 2005) or Kratochwill protocol (UCL Educational Psychology Literature Review Coding Protocol, adapted from the

APA Task Force Coding Protocol, 2003), alongside Gough's (2007) Weight of Evidence Framework. All of the Precision Teaching interventions were based on the test- retest model; where scores were taken from assessments pre and post intervention, with some taking mid-intervention and follow up scores.

All of the studies identified that Precision Teaching had a positive effect on the word reading skills of students, with the range of effects varying from small to large. Methodological limitations are explored and further research has been suggested as necessary to strengthen the evidence based of this intervention.

Section 2: Introduction

What is Precision Teaching?

Precision Teaching is a school based intervention aimed at accelerating student progress in specific areas, dependent on which area has been highlighted as an area of difficulty for that individual student. Precision Teaching is not a teaching method per se (Binder, 1988; Pennypacker, Koenig, & Lindsley, 1972) , but is best recognised as a measurement tool of learning (West , Young & Spooner, 1995) that includes access to repeated practices, error-correction and timed drills . The results from the Precision Teaching intervention are recorded on the Standard Celeration Charts (SCC) *.

*The SCC is a precise measurement system which provides the student and the teacher with an overall picture of the student's progress. Calendar days are recorded on the x-axis and count-per-minute on the y-axis.

Precision Teaching can sit alongside any curriculum or learning area with ease (Kubina & Yurich, 2012) and it acts as an analytical tool which can help decipher if an instructional classroom based teaching method is achieving success (Chiesa & Robertson, 2000).

Precision Teaching is a student-centred approach (Boyce, 2003) and it focused on an area of need that is individual to the student, for example word reading. It can be an effective intervention when students lack the confidence or motivation to complete or try tasks, or if students appear to know something but then struggle to retain it over time, and for students who are very slow at producing work.

The Precision Teaching approach provides a framework whereby the student has access to daily sessions of teaching, frequency building, monitoring and the opportunity to evaluate progress on a regular basis (Lindsley, 1992). As and when the student becomes competent in the desired domain, the learning programme is then adapted to allow the student to progress onto more challenging tasks.

Due to the flexibility within this approach, and its student-centredness, it is an intervention that can be used with students of all ages and abilities (Gallagher, Bones & Lambe, 2016) and is an approach that can be implemented on a whole class level, although it is most effective when carried out with individual children (Chiesa & Robertson, 2000). If the student is making progress within the Precision Teaching sessions then the programme is deemed as appropriate, but if progress is not being seen, then the deficit is with the programme itself and not the student (Fan-Yu & Kubina, 2004).

One of the unique characteristics of the Precision Teaching approach is the daily charting (Standard Celeration Charts) of student progress based on observable behaviours within the Precision Teaching sessions (Gallagher et al., 2016). It has the benefit of an in-built monitoring system, which provides the student and teacher with instant feedback on the student's progress; this acts as a form of evaluation in itself. Feedback is an important aspect of the learning process and is crucial in improving learning outcomes for students (Hattie, 2009).

Solity and Bull (1987) devised "five basic steps" within the Precision Teaching intervention. These include:

1. Specifying the students task in observable, measureable terms with the use of probes*
2. Recording the students' progress on a daily basis
3. Charting the students' performance on a daily basis
4. Recording the teaching approach in relation to the students' progress
5. Analysing the data to determine whether:
 - Progress is satisfactory
 - Changes are needed in the teaching approach in order to maintain or accelerate learning.

*Probes are used to monitor the extent to which the student is accurate and fluent in that skill area. Probes can take many forms, from word lists to mathematical sums. These probes are available commercially however they can be created easily; this is useful as it can be tailored to the individual student.

These basic steps devised by Solity and Bull (1987) should be followed by those carrying out a Precision Teaching intervention.

Successful examples of Precision Teaching in practice have been evidenced in longitudinal studies in the USA namely, The Precision Teaching Project in the Great Falls, Montana (Beck & Clement, 1991) Morningside Academy (Binder, 1990; Lindsley, 1992).

Evidence from these studies showed the effectiveness of the Precision Teaching approach in improving pupil performance across the curriculum.

Doughty, Chase, and Shields (2004) carried out a literature review on some of the principles of Precision Teaching including 'fluency', 'endurance' and 'repeated readings' although this review was not specific to word reading. Only one of the six studies in this review (Kessissoglou & Farrell, 1995) was included in this 2004 review. Since 2004 the UK and Ireland research base has increased, hence the need for a review of the effectiveness of this approach in UK and Ireland specifically.

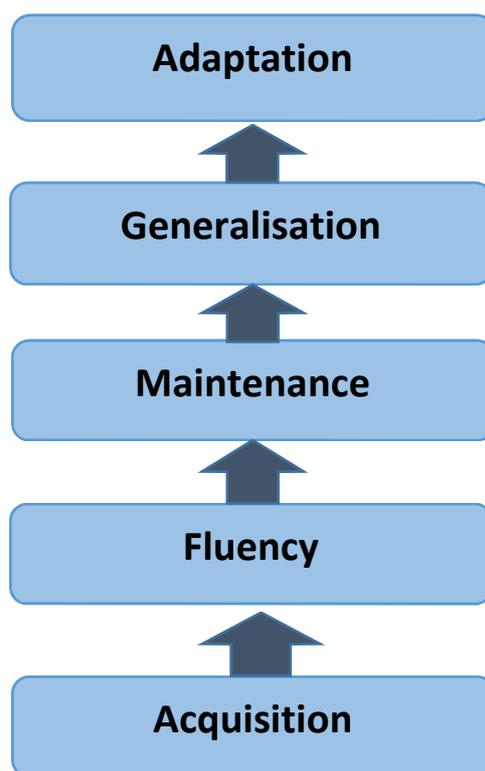
Psychological basis of Precision Teaching

Precision Teaching's psychological theory is primarily rooted in behavioural analysis and was first coined by Ogden Lindsley in the 1960's in the USA. Lindsley was a student of B.F Skinner, and recognised the potential value of recreating the techniques used by Skinner, namely free operant conditioning and analysing rates of responding within the classroom. Pavlov's use of frequency to measure observable behaviour also underpins the Precision Teaching approach (Potts, Eshleman & Cooper, 1993).

Precision Teaching supports the Vygotskian (1978) theory of the ‘Zone of Proximal Development’ (ZPD) which encourages educators to be very specific about what they are teaching students to ensure that the gap can be bridged between what the student does and doesn’t know, and what they can potentially achieve with the guidance, encouragement and scaffolding from an educator. Wood, Bruner and Ross (1976) affirms that students are most likely to experience success when challenging tasks are scaffolded and modelled.

The principles that underpin Precision Teaching, in terms of the learning process sit alongside Haring and Easton’s (1978) ‘Hierarchy of Learning’ with ease (*Diagram 1*).

Diagram 1:



The diagram shows that as the student moves through the stages between *Acquisition to Adaptation* the student learns new skills, building up accuracy, fluency

and it is only then that they are able to maintain the new skills, before being able to generalise these new skills within new conditions; at first with instruction and then independently.

Rationale and relevance to Educational Psychology practice

Research indicates that those who experience difficulties with literacy are more likely to require special educational needs provision, to truant from educational settings, and to be excluded (Dugdale & Clark, 2008). The National Literacy Strategy (1998) highlighted the importance of evidenced based, targeted interventions to those students who exhibit the poorest attainment. This included introducing personalised learning programmes and increasing opportunity to specific literacy interventions. Although the teaching of literacy within the '*Literacy Hour*' improved following the introduction of the National Literacy Strategy (1998), reviews found this was not informing teaching sufficiently (Ofsted, 2005) and that more needed to be done.

The recent figures from the Organisation for Economic Co-operation and Development (OECD, 2016) showed that England had amongst the lowest literacy levels in all of the OECD countries, lagging behind countries such as Poland and Estonia. The Department for Education (2001) states that Precision Teaching is well placed to support in the development and modification of teaching programmes, and to evaluate the effectiveness to these programmes. Moreover, The Rose Report (2009) identified Precision teaching as an effective intervention for improving the literacy skills of students, emphasising that interventions were most successful when they were "focussed on the right children through careful assessment, regular

updating and tracking of progress” (Rose 2009, p. 45). Even a few minutes per day of timed practice can often eliminate what were previously considered irremediable learning problems (Binder, 1988).

More recently, the Department for Education (2015) highlights the importance within education of having a culture of high aspirations and stretching targets for students, alongside students having choice and control over their learning. These principles are at the core of the Precision Teaching approach and provide further arguments for Educational Psychologists recommending and schools implementing a Precision Teaching approach.

Doughty et al. (2004) argue that Precision Teaching has made a significant impact on the education system in the United States, however the evidence base in the United Kingdom and Ireland has been sporadic (Gallagher et al., 2016). In spite of the limited research base in the UK, it continues to be recommended as an effective intervention by Educational psychologists in the UK (Chiesa & Robertson 2000; Kessissoglou & Farrell 1995).

Review question

How effective is Precision Teaching in improving the word reading skills of school aged students in the United Kingdom and Ireland?

Section 3: Critical review of the evidence base

Literature search

A comprehensive literature search was conducted in December 2015 using four electronic databases; PsychINFO, ERIC, Medline and Web of Science. Using a multi-field search the following search terms were entered (*Table 1*).

Table 1: Database search terms

DATABASE	SEARCH TERMS
PsychINFO	Contains 'Precision Teaching' AND 'Literacy'
ERIC	Contains 'Precision Teaching' AND 'Literacy'
Medline	'Precision Teaching'
Web of Science	'Precision Teaching'

Following the database search a further search of the Journal of Precision Teaching & Celeration and an ancestral search was conducted. *Figure 1* provides a flowchart to illustrate the literature search. *Table 2* shows the inclusion and exclusion criteria specified within this search. *Table 3* shows the studies included within this review and those excluded after an inspection of the full article can be found in *Appendix 1*. The summary of all the studies in *Table 3* can be found in *Appendix 2*.

Figure 1: Flow diagram of literature search

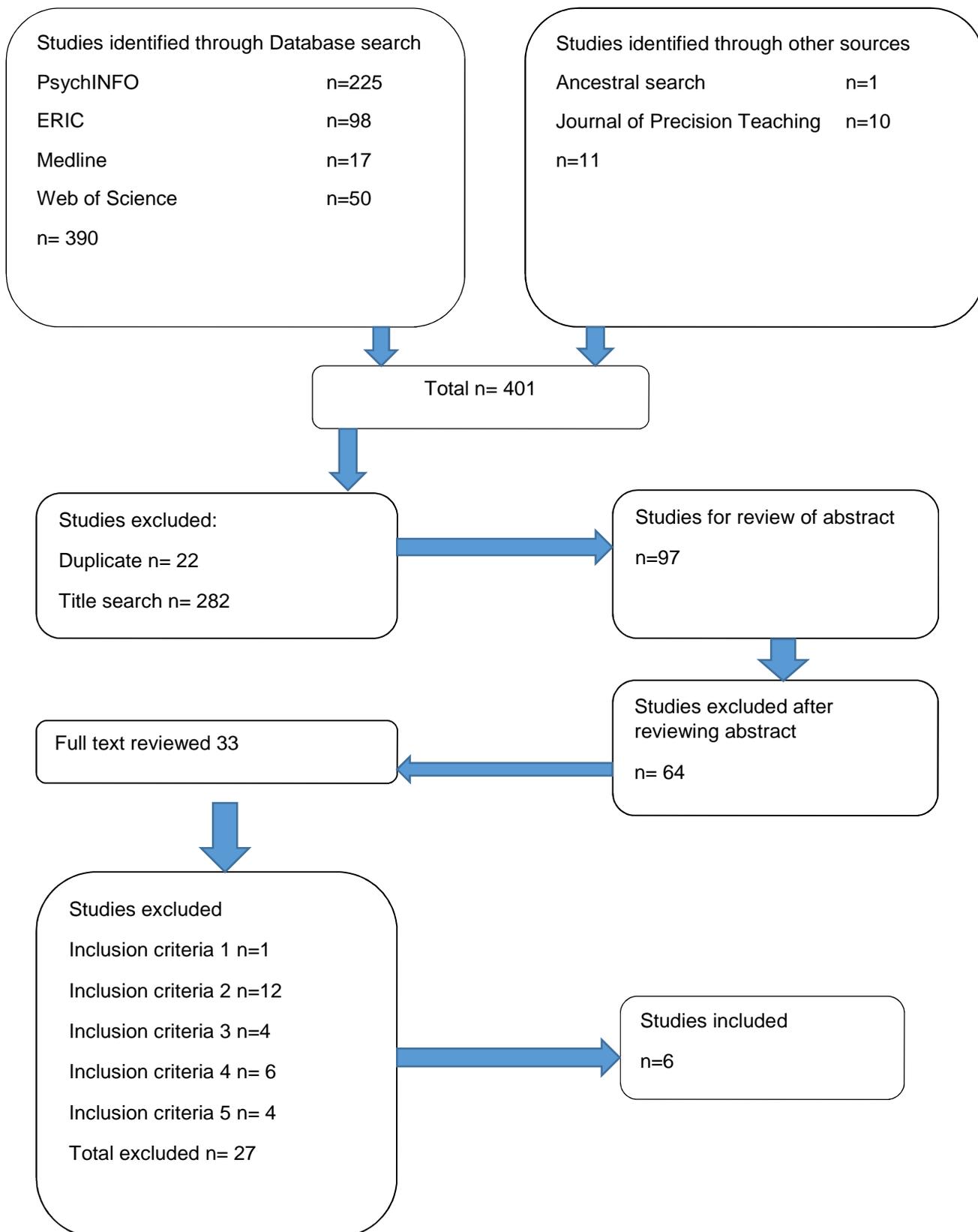


Table 2: Inclusion and exclusion criteria

		Inclusion Criteria	Exclusion Criteria	Justification
1	Type of Publication	Must be in a peer reviewed journal.	Not in a peer reviewed journal.	Peer reviewed studies have undergone a level of scrutiny.
2	Language and Country of Origin	Studies written in English and undertaken in the UK and Ireland.	Studies not written in English or was not undertaken in UK and Ireland.	Currently there is a limited research base in the UK and Ireland.
3	Setting	Primary and secondary schools	Not in a school i.e. an afterschool club, nursery or University.	To consider the implication of the intervention on students in schools only.
4	Intervention	Precision Teaching.	Not 'Rate Building', 'Fluency Training' 'speed reading' or 'Direct Instruction'	The reviewer would not be able to discern whether effects were simply from a Precision Teaching intervention.
5	Type of design	Group based or Single case study design.	Not group based or single case design.	The review compared the effectiveness of the intervention either between or within participants.
6	Participants	Participants are of school age.	Participants are not of school age, i.e. Nursery or University age.	To consider the implication of the intervention on statutory school aged children
7	Measures	The study must have pre and post intervention data.	The study does not have pre or post intervention data.	This helps to establish a direct measurement of the effectiveness of the intervention.

Table 3: Selected studies for inclusion in review

Included studies
<ol style="list-style-type: none"> 1. Roberts, W., & Norwich, B. (2010). Using precision teaching to enhance the word reading skills and academic self-concept of secondary school students: a role for professional educational psychologists. <i>Educational Psychology in Practice</i>, 26(3), 279–298. 2. Hughes, J. C., Beverley, M., & Whitehead, J. (2007). Using Precision Teaching to increase the fluency of word reading with problem readers. <i>European Journal of Behavior Analysis</i>, 8(2), 221–238. 3. Griffin, C. P., & Murtagh, L. (2015). Increasing the sight vocabulary and reading fluency of children requiring reading support: the use of a Precision Teaching approach. <i>Educational Psychology in Practice</i>, 31(2), 186–209. 4. Downer, A. C. (2007). The National Literacy Strategy Sight Recognition Programme implemented by teaching assistants: A precision teaching approach. <i>Educational Psychology in Practice</i>, 23(2), 129–143. http://doi.org/10.1080/02667360701320820 5. Lambe, D., Murphy, C. & Kelly, M.E. (2015). The Impact of a Precision Teaching Intervention on the Reading Fluency of Typically Developing Children. <i>Behavioural Interventions</i>. 30: 364-377. 6. Kessissoglou, S. & Farrell, P. (1995). Whatever happened to Precision Teaching? <i>British Journal of Special Education</i>. Vol 22 (2).

Critical Review of studies

The six studies that were selected for review were coded using the Quality Indicators: Within Single-Subject Research by Horner et al. (2005) and UCL Educational Psychology Literature Review Coding Protocol, adapted from the APA Task Force Coding Protocol by Kratochwill (2003). These studies were then appraised for quality and relevance based on the 'Weight of Evidence Framework' (Gough, 2007) which includes making judgements about the quality and relevance of that evidence. This was used to determine how each study was significant in

answering the specific review question. Gough’s (2007) framework helps the reviewer to come to a more objective stance as to how the review question is valid to the individually selected studies. For those studies which were single case design (Downer, 2007; Hughes et al., 2007; Lambe et al., 2015) , these were rated against Horner et al. (2005) for one of the elements in WOE A. The Harden and Gough (2012) framework (*Table 4*) includes information regarding how the overall Weight of Evidence is determined. *Table 5* shows the overall rating of evidence for each of the six chosen studies. The full Weight of Evidence criteria can be found in *Appendix 3*.

Table 4: Framework for Weight of Evidence (Harden & Gough, 2012)

Weight of evidence A	Weight of evidence B	Weight of evidence C	Weight of evidence D
Quality of execution of the study in relation to quality standards for studies of that type (<i>Methodological Quality</i>)	Appropriateness of research design for addressing Review Question (<i>Methodological Relevance</i>)	Appropriateness of focus of study to Review Question (<i>Topic Relevance</i>)	Considering A, B & C to rate the overall degree to which the study contributes in answering the Review Question (<i>Overall weight of evidence</i>)

Table 5: Overall Weight of Evidence for each study

Studies	Weighting score			Overall Weight of evidence D
	Methodological Quality A	Methodological Relevance B	Topic Relevance C	
Roberts & Norwich (2010)	Low (1)	Medium (2)	Medium (2)	Medium (1.6)
Hughes et al. (2007)	Low (1)	Low (1)	Medium (2)	Low (1.3)
Lambe et al. (2015)	Medium (1.7)	Low (1)	Medium (2)	Medium (1.6)
Downer (2007)	Medium (2)	Medium (2)	Medium (2)	Medium (2)
Griffin & Murtagh (2015)	Medium (2)	Medium (2)	Low (1)	Medium (1.6)
Kessissoglou & Farrell (1995)	Low (1)	Low (1)	Low (1)	Low (1)

Participants

The studies chosen for review were conducted in either the United Kingdom (Downer, 2007; Hughes et al., 2007; Kessissoglou & Farrell, 1995; Roberts & Norwich, 2010) or Ireland (Griffin & Murtagh, 2015; Lambe et al., 2015). A total of 196 participants were included in this review with ages ranging from 4 years – 16 years. Within the Robert and Norwich (2010), Downer (2007), Griffin and Murtagh (2015), Hughes et al. (2007) and Kessissoglou and Farrell (1995) studies, participants were chosen due to their reading levels being below their chronological age. Lambe et al. (2015) participants were not reported to have reading difficulties or learning needs on entry to the study. Downer (2007) was the only study to report and consider gender differences within its findings, although no significant difference was found.

Design

Of the six studies, five used a quantitative only approach, with Griffin and Murtagh (2015) using a mixed methods approach (semi- structured interviews with teachers and students). All studies took pre and post measures of the participants before and after the intervention and three of the studies took follow up measures (Hughes et al., 2007; Kessissoglou & Farrell, 1995; Roberts & Norwich, 2010). The research designs were varied; Hughes et al. (2007), Downer (2007) and Lambe et al. (2015) used a single subject design, with the latter two studies also using Multiple Baseline Design (MBD). It must be noted however, that although Hughes et al. (2007) was described as a single subject design by the research authors, there was in fact two participants (out of the total of seven) who had treatment as usual (TAU). Therefore it could be argued that it employed a group based design. For the purposes of this review however, Hughes et al. (2007) will be referred to as a single subject design. Robert and Norwich (2010), Griffin and Murtagh (2015) and Kessissoglou and Farrell (1995) used a quasi-experimental group design.

Griffin and Murtagh (2015) and Kessissoglou and Farrell (1995) had control groups and within these studies control and intervention groups were randomly assigned; this was reflected in the WOE B and these studies received a higher rating. Trained Teaching Assistants (TA's) delivered the interventions in all of the studies but one (Kessissoglou & Farrell, 1995), where one of the authors (Kessissoglou) implemented the intervention.

Robert and Norwich (2010) used a waitlist control group, which is deemed to be more ethical than a no intervention group, as the intervention which the researcher

believes to be effective is not being withheld from participants (Barker, Pistrang & Elliot, 2002). Within the single case design studies, the baseline condition acted as a within-participant control, similar to how a control group is used within a group design study (Horner et al., 2005), therefore these studies were not disadvantaged due to their methodological stance as seen in WOE A. However Horner et al. (2005) also state that within MBD there should typically be at least 5 measuring points; neither of the two studies had five measuring points which was reflected in the WOE A and B. Consequently, it is worth considering whether the students in the single case design studies would have improved their word reading skills regardless of the intervention due to natural maturational factors, or with normal curriculum teaching. This reflects one of the potential threats to the internal validity of pre-test and post-test designs (Cook & Campbell, 1979 as cited in Barker, Pistrang & Elliot, 2002).

Quality of the Intervention

The quality of the interventions were reviewed using the following themes: frequency, duration and manualisation of the interventions, treatment fidelity and outcome measures.

Frequency, duration and manualisation of the intervention

Intervention periods varied considerably between studies. Downer's (2007) intervention period ranged from 2-26 weeks with 22 being the average intervention period. The remaining studies intervention periods ranged from 5- 10 weeks. Within the Robert and Norwich (2010), Downer (2007), Kessissoglou and Farrell (1995) and Griffin and Murtagh (2015) studies the interventions took place 5 days per week.

Within the Hughes et al. (2007) and Lambe et al. (2015) studies the intervention took place 3-4 times per week. Ideally Precision Teaching should be implemented on a daily basis, so that fluency can be maintained, however there can be flexibility within this intervention. The frequency and duration of the intervention was rated in WOE C, with those adhering strictly to Precision Teaching principles (e.g. daily practice) receiving a higher rating.

The difference in intervention periods across studies should be considered with regards to the overall effect of the intervention. The participants in the Hughes et al. (2007) study received the intervention for the longest duration of time (approximately 700 minutes over the 10 week intervention), however the participants in the Kessissoglou and Farrell (1995) study only received the intervention for 375 minutes in total. Highest effect sizes were shown for the Hughes et al. (2007) study, posing the question whether the duration of the intervention has a direct effect on the effectiveness of the intervention. Effect sizes were not able to be calculated for the Kessissoglou and Farrell (1995) study due to insufficient data being presented within the paper. One of the unique features of Precision Teaching intervention is the use of a Standard Celeration Chart (SCC). Downer (2007), Griffin and Murtagh (2015), Hughes et al. (2007), Lambe et al. (2015) used the SCC to monitor and record student progress. Roberts and Norwich (2010) noted that they monitored student progress on standard record forms, but it is unsure whether this was the SCC or a form unique to their study. It is unsure what form of recording and monitoring Kessissoglou and Farrell (1995) used. All of the studies explained, in enough detail, the procedures and measures that were used, which means that the studies could be replicated. However, due to the amount of variability (in terms of the length of intervention periods, and the measures used to assess the student's pre and post

intervention) it is difficult to directly compare the studies because of these differences.

Fidelity of intervention

Fidelity of the implementation of an intervention is an important aspect to consider, and is an essential component of WOE. C. Lambe et al. (2015) reported an inter-observer reliability mean rate of 99.5% and Hughes et al. (2007) reported an inter-observer reliability rate of 97.5%. Roberts and Norwich (2010) took steps to monitor the fidelity by observing the TA's who were implementing the intervention to ensure that they were adhering to the agreed format. Inter-observer reliability is useful as it indicates how consistent the intervention was implemented across students. The remaining studies did not provide information regarding the fidelity of the implementation of the intervention.

Downer (2007), Roberts and Norwich (2010), Lambe et al. (2015), Griffin and Murtagh (2015) all noted that Teaching Assistants had had prior training before implementing the intervention. Hughes et al. (2007) did not note this, and within the Kessissoglou and Farrell (1995) study, one of the authors Kessissoglou was also the person implementing the intervention.

Outcome measures

Outcome measures differed greatly between studies and included; Say All Fast a Minute Every Day Shuffled (SAFMEDS, cited in Lambe et al., 2015), Gray Oral Reading Test (GORT; Bryant & Wiederholt, 2011), Reading Accuracy Measure (RAM; cited in Roberts & Norwich, 2010), Group Reading Test II (NFER-Nelson, 2000) the vocabulary scale of Middle Years Information System (MidYis; CEM, 2007),

Dolch story (cited in Lambe et al., 2015), Salford Sentence Reading test (Hodder & Stoughton, 2000), The Neale Analysis of Reading Ability (NARA; Neale, 1999) and the Special Needs Information Pack (SNIP-PT-Pack; Smart & Smart, 2007). None of the studies reported reliability scores for the measures used in assessing word reading levels which was reflected in with lower WOE A ratings.

Only two of the six studies used appropriate methods of analysis for results. Lambe et al. (2015) used Partial eta squared and Roberts and Norwich (2010) used Kolmogorov-Smirnoff Z and recorded the effect size using correlation (r). As the remaining studies did not calculate effect sizes, these were calculated using Pre-test post-test Standardised Mean Difference referred to as PPSMD (Becker, 1988) to establish whether the interventions were effective.

Kessissoglou and Farrell (1995) did not provide enough information to calculate PPSMD therefore this study received a lower WOE A. *Table 6* shows the outcomes and effect sizes for the studies and *Table 7* gives a description of small, medium and large effect sizes.

Sample sizes within the studies varied between 77 and 7 participants. Cohen (1992) states that researchers must know the sample size required to obtain the desired power and effect size. Cohen (1992) advises that in order to detect a large effect size it is necessary for group sizes to consist of at least 26 participants, 64 participants to detect a medium effect and 393 participants to detect a small effect (when comparing means at the 0.5 significance level and when the power level is set at 0.8). Cohen (1992) also states that when undertaking statistical tests for two or more groups that the sample size that is defined above, is for each group. By Cohen's (1992) interpretation the three group based designs within this review (Downer et al., 2007; Griffin & Murtagh; Roberts & Norwich, 2010) were all

underpowered. One must therefore be cautious when interpreting these effect sizes, as studies which lack power are more likely to produce non-significant results and cannot be generalised across the population. This has been reflected in WOE A, and these studies were assigned lower ratings accordingly. Furthermore, three of the studies in this review employed a single case design (sometimes called small n designs). Within single case designs the individual participants serve as their own control, and sample size can be as low as one participant. Consequently, power and sample size is not relevant to this type of design. To reflect this a separate criteria for WOE A and B was utilised (based on Horner et al., 2005) to establish methodological quality. This ensured that the single case designs in this review were not disadvantaged due to their methodological stance.

Findings

In all but one of the studies, effect sizes were either given (Griffin & Murtagh, 2015; Robert & Norwich, 2010) or were calculated using Pre-Post Standardised Mean Difference (PPSMD) by the reviewer (Downer, 2007; Hughes et al., 2007; Lambe et al., 2015). Kessissoglou and Farrell (1995) did not provide effect sizes nor was enough information given for these to be calculated. However, the data provided within the study indicated that the participants receiving the intervention increased their word accuracy skills by seven months and their reading comprehension by eighteen months (after the five week intervention). The remaining studies recorded large effect sizes providing further evidence to support the use of Precision Teaching as an effective intervention.

Table 6: Effect sizes

Study	Outcome measure	Participants	Effect size	Overall Weight of Evidence
Roberts & Norwich (2010)	Reading Accuracy Measure (RAM)	n=77 Total n=77	Cohort 1 Group A (PT) and Group B (UTA) r=.14 (pre to mid measure) Group A (UTA) and Group B (PT) r=.10 (mid to post) Cohort 2 Group C (PT) and Group D(UTA) r=.21 (pre-mid) Group C (UTA) and Group D (PT) r=.15 (Kolmogorov-Smirnov Z) Small	Medium
Hughes et al. (2007)	GRT II & Midyis	n=5 n=2 Total n=7	Intervention PPSMD=6.8 Large TAU PPSMD=0.95	Medium
Lambe et al. (2015)	SAFMEDS	n=7	Baseline- post PPSMD=3.98 Baseline-follow up PPSMD=3.74 Large	Medium
	Dolch story	n=7 Total n=7	Baseline- post PPSMD=2.76 Baseline-follow up PPSMD=2.35 Large	
Downer (2007)	Word list	n=13	Reading score Group 1 (Years 1 & 2) PPSMD= 7.08 Large	Medium
	Word List	n=19	Large Group 2 (Years 3,4,5,6)	

	Salford Reading Test	n=15	PPSMD=1.17 Large	
		Total n=47	Reading age (months) Group 3 (Years 7 & 8) PPSMD=1.4 Large	
Griffin & Murtagh (2015)	SNIPP-T & Gort 4	n=40	Reading accuracy (eta squared =.69) Reading fluency (eta squared=.64) Reading comprehension (eta squared =.55) Overall reading ability (eta squared =.72) Large	Medium
Kessissoglou & Farrell (1995)	NARA	Total n=40 n=12	No effect sizes provided and not enough information was given to calculate these.	Low
		Total n=12		

Table 7: Description of small, medium and large effect sizes

Type of effect size	Small	Medium	Large
(Cohen, 1998)	0.2	0.5	0.8
Partial eta squared (Cohen, 1988)	0.01	0.06	0.14
Kolmogorov-Smirnov Z (reported as correlational (r))	0.3	0.5	0.8

Consideration should be given to the aspect of the age of participants, as Roberts and Norwich (2010) was the only study that was solely undertaken within secondary schools, although Downer (2007) did have some participants of secondary school age; the youngest students in this study (primary age pupils) showed the largest effect sizes. Despite the reassuring medium- large effect sizes seen in four of the six studies, these results should be interpreted with caution due to small sample sizes. Robert and Norwich (2010) showed the smallest effect sizes of all the studies and this study had the largest sample size (n=77). The studies with smaller sample sizes reported larger effect sizes and again this must be viewed with caution due to the lack of generalisability across larger populations. However, it is not surprising that those studies who showed larger effect sizes had smaller sample sizes (n) as the likelihood of seeing a large effect is greater the smaller the sample size (Slavin & Smith, 2009).

Furthermore, Downer (2007) reported the largest effect size (PPSMD=7.08) but this study had the longest intervention period (average of 22 weeks amongst participants) therefore consideration must be given to the correlation between the length of the intervention and its effectiveness.

Due to the nature of Precision Teaching (repeated measures), the participants would have been exposed repeatedly to a variation of the same words throughout the intervention , therefore it could be expected that the more exposure a student has to the words the more likely they will be able to recite them with accuracy and fluency post intervention.

Section 4: Conclusions and Recommendations

The aim of this review was to evaluate the effectiveness of a Precision Teaching intervention in school aged students. Six studies were identified that met the inclusion criteria and have been discussed throughout this review paper. All but one study (Roberts & Norwich, 2010) showed large effect sizes from the implementation for the Precision Teaching interventions. Kessissoglou and Farrell (1995) and Hughes et al. (2007) received a 'Low' rating using Gough's (2007) Weight of Evidence Framework. The remaining four studies received 'Medium' ratings using the same framework. Although this review indicates some positive evidence to suggest that Precision Teaching is an effective intervention for reading difficulties, further research is needed comparing the use of Precision Teaching against other reading interventions.

One of the main issues was the range of outcome measures that were used when implementing the Precision Teaching approach and consideration given to whether the same students would receive the same results if different outcome measures were used; none of the studies reported the reliability of the outcome measures.

When implementing a Precision Teaching intervention it would be important for educators to be specific about the outcome measures, and to use measures with high reliability. Also the length of intervention periods varied greatly within these studies, and therefore it would be important for educators to monitor, review and evaluate this intervention at regular periods. Precision Teaching is an intervention delivered on a one: one basis and therefore educators should be mindful of the

potential benefits of this intervention against others that could perhaps be carried out on a group level.

A number of factors have made direct comparison between the studies problematic including the variation in methodological approaches, the measurement tools used, the varying levels of intensity, and the duration of the interventions.

Recommendations

- More research needs to be carried out and published within the UK and Ireland to add the evidence base of the effectiveness of Precision Teaching
- Stronger evidence of efficacy and maintenance of skills post intervention is necessary to show the longer term effects of the intervention.
- Studies should provide standardised effect sizes so that comparisons can be made between and within this interventions or if using MBD then at least 5 measurement points should be included.
- Further research needs to be carried out on diverse groups of students, and those with different levels of abilities. These studies should be large enough to give confidence that the results will hold in different populations.

References

- Barker, C., Pistrang, N., & Elliott, R. (2002). *Research methods in clinical psychology: An introduction for students and practitioners*. (2nd ed) Chichester: John Wiley & Sons LTD.
- Beck, R., & Clement, R. (1991). The Great Falls Precision Teaching Project : An Historical Examination, (1).
- Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257–278.
- Binder, C. (1988). Precision Teaching : Measuring and attaining exemplary Precision Teaching : Measuring and Attaining Exemplary Academic Achievement, 10, 12–15.
- Binder, C. (1990). Binder, C. Precision teaching and curriculum based measurement. *Review Literature And Arts Of The Americas*, 7(2), 33–35.
- Binder, C., & Watkins, C. L. (1990). Precision Teaching and Direct Instruction : Measurably superior instructional technology in schools . Performance Improvement Quarterly , 3 (4), 74-96 ., 3.
- Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst*, 19(2), 163-197.
- Boyce, T. E. (2003). Moving from precision teaching to precision measurement: We need to just DO IT. *European Journal of Behavior Analysis*, 4, 59-63.
- CEM. (2007). *Middle Years Information System (MidYIS)*.
- 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. <http://doi.org/10.1080/713666088>.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. New Jersey: Lawrence Erlbaum Associates, Inc. Publishers. pp 283-286.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155.
- Department for Education and Employment (1998). *The national literacy strategy. Framework for teaching*. London: DfES.
- Department for Education (DfE). (2001). *Special Educational Needs Code of Practice*. London: HMSO.
- Department for Education (2015). *Special educational needs and disability code of practice: 0 to 25 years*. London: Department for Education.
- Doughty, S. S., Chase, P. N., & Shields, E. M. O. (2004). Effects of Rate Building on Fluent Performance : A Review and Commentary, 1(1), 7–23.

- Downer, A. C. (2007). The National Literacy Strategy Sight Recognition Programme implemented by teaching assistants: A precision teaching approach. *Educational Psychology in Practice*, 23(2), 129–143. <http://doi.org/10.1080/02667360701320820>.
- Dugdale, G., & Clark, C. (2008). *Literacy changes lives: An advocacy resource*: National Literacy Trust London.
- Fan-Yu Lin and Kubina R.M. (2004) Learning Channels and Verbal Behavior, *The Behavior Analyst Today*, 5 (1), 1-14.
- Gallagher, E., Bones, R., & Lambe, J. (2016). Precision teaching and education : Is fluency the missing link between success and failure ?. *Irish Educational Studies* Vol. 25, No. 1, pp. 93-105.
- Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213–228.
- Griffin, C. P., & Murtagh, L. (2015). Increasing the sight vocabulary and reading fluency of children requiring reading support: the use of a Precision Teaching approach. *Educational Psychology in Practice*, 31(2), 186–209.
- Harden, A., & Gough, D. (2012). Quality and Relevance Appraisal. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An Introduction to Systematic Reviews* (pp. 153–178). London: Sage.
- Haring, N.G., Lovitt, T.C., Eaton, M.D., & Hansen, C.L. (1978). *The fourth R: Research in the classroom*. Columbus, OH: Charles E. Merrill Publishing Co.
- Hodder and Stoughton (2000) Salford Sentence Reading test.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165–179.
- Hughes, J. C., Beverley, M., & Whitehead, J. (2007). Using Precision Teaching to increase the fluency of word reading with problem readers. *European Journal of Behavior Analysis*, 8(2), 221–238.
- Kessissoglou, S. & Farrell, P. (1995) Whatever happened to Precision Teaching? *British Journal of Special Education*. Vol 22 (2).
- Kratochwill, T. R. (2003). Evidence-Based Practice: Promoting Evidence-Based Interventions in School Psychology. *School Psychology Quarterly*, 18(4), 389-408.
- Kubina, R. M., & Yurich, K. K. L. (2012). *The Precision Teaching Book*, 1–18.
- Lambe, D., Murphy, C. & Kelly, M.E. (2015) The Impact of a Precision Teaching Intervention on the Reading Fluency of Typically Developing Children. *Behavioural Interventions*. 30: 364-377.
- Lindsley, O. R. (1992). Precision teaching: Discoveries and effects. *Journal of*

Applied Behavior Analysis, 25(1), 51–57. <http://doi.org/10.1901/jaba.1992.25-51>.

- Neale, M. D. (1999). *Neale analysis of reading ability: Manual*. ACER Press, Australian Council for Educational Research Limited, Australia.
- NFER-Nelson. (2000). *Group Reading Test II (6-14)*. Windsor: Author.
- Organisation for Economic Co-operation and Development (2016) Building skills for all: A review for England.
- Ofsted. (2005). *The National Literacy and Numeracy Strategies and the National Curriculum*. London: Office for Standards in Education (Ofsted).
- Pennypacker, H. S., Koenig, C. H., & Lindsley, O. R. (1972). *The handbook of the standard behavior chart*. Kansas City, KS: Behavior Research Co.
- Potts, L., Eshleman, J. W., & Cooper, J. (1993). Ogden R . Lindsley and the Historical Development of Precision Teaching, 2(2), 177–189.
- Roberts, W., & Norwich, B. (2010). Using precision teaching to enhance the word reading skills and academic self-concept of secondary school students: a role for professional educational psychologists. *Educational Psychology in Practice*, 26(3), 279–298. <http://doi.org/10.1080/02667363.2010.495215>
- Rose, S. J., & June, F. (2009). Identifying and Teaching Children and Young People with Dyslexia and Literacy Difficulties An independent report from Sir Jim Rose to the. *Secretary*, (June).
- Slavin, R., & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational Evaluation and Policy Analysis*, 31(4), 500-506.
- Smart, C., & Smart, P. (2007). SNIP Precision Teaching pack. Retrieved from http://www.snip-newsletter.co.uk/pdfs/downloads/precision_teaching.pdf
- Solity. J., & Bull, S. (1987). *Special needs: Bridging the curriculum gap*. Milton Keynes: Open University Press.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- West, R. P., Young, R. & Spooner, F. (1995) Precision teaching: an introduction, *Journal of Precision Teaching*, 12, 2_8.
- Wiederholt, J. L., & Bryant, B. R. (2001). *Gray oral reading test (4th ed.)*. Austin, TX: Pro-Ed.
- Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Child Psychiatry*, 17, 89–100.

Appendix 1: List of excluded studies

Excluded Studies	Reason for exclusion
1. Brent, G. (1978) Urban Applications of Precision teaching to Reading. <i>Reading Improvement</i> , v15 n3 p190-202	Criteria 2
2. Kubina, R M., Commons, M L., Heckart, B (2009) Using Precision Teaching with Direct Instruction in a Summer School Program, <i>Journal of Direct Instruction</i> 9.1: 1-12.	Criteria 2, 3 & 4
3. Spencer, M L., Henderson, J, C. (1974-75) Sante Fe School Precision Teaching Program, Evaluation Report 1974-75. Telegraph Development Co, Berkeley, CA.	Criteria 2
4. Kent, J., Street, E. M. (2013) Response to Intervention and precision teaching: Creating synergy in the classroom. Guildford Press. US.	Criteria 2
5. Selfridge, K A., Kostewicz D E. (2011) Reading interventions for students with learning disabilities. <i>Journal of Precision Teaching & Celeration</i> . Vol.27, pp19-24.	Criteria 2
6. Cavallini, F., Berardo, F., Perini, S (2010) Mental retardation and reading rate: Effects of precision teaching. <i>Life Span and Disability</i> . Vol 13 (1), pp.87-101.	Criteria 2
7. Ascah, K. (2009) Regina’s reading program and progress. <i>Journal of Precision Teaching & Celeration</i> . Vol.25 2009, pp.35-39.	Criteria 2
8. Syrek, A L., Hixson, M D., Jacob, S., Morgan, S. (2007) A Comparison of error correction procedures on word reading. <i>Journal of Precision Teaching & Celeration</i> . Vol.23, pp. 2-13.	Criteria 2
9. Sulgrove, M K., McLaughlin, T.F. (2004) The Effects of an Additional Timed Reading on Reading Rate. <i>Journal of Precision Teaching & Celeration</i> . Vol.20 (1), pp. 9-16.	Criteria 2
10. Black, A, R., Somers, M., Doolittle, F., Unterman, R., Grossman, J B et al (2009) The Evaluations of Enhanced Academic Instruction in After-School Programs: The Final Report. NCEE.	Criteria 3
11. Fabrizio, M A., Schimer, K., Ferris, K. (2002) Tracking Curricular Progress with Precision, <i>Journal of Precision Teaching & Celeration</i> . Vol.18 (2), pp. 78-79.	Criteria 2
12. Freeze, R., Cook, P (2005) Learning to read against all odds: Using precision reading to enhance literacy in students with cognitive impairments, extreme academic deficits, and severe social, emotional and	Criteria 3

- psychiatric problems. *Exceptionality Education Canada*. Vol.15 (1), pp. 79-109.
13. Johnson, K. (1997). Morningside Academy, 7(1), 31–35. Criteria 3
 14. Peterson, S K. (1990) Using the Language Experience with Precision. *Teaching Exceptional Children*. 22.3: 28-31. Criteria 4
 15. Mountseven, J. (1990) Speed Reading: A Technique for Developing Fluent Readers. *Teaching Exceptional Children* 22.3 : 69-71. Criteria 4
 16. Malanga, P. (2003) Using Repeated Readings and Error Correction to Build Fluency with At Risk Elementary Students. *Journal of Precision Teaching & Celeration*. Vol.19 (2), pp. 19-27. Criteria 4
 17. Legault, A., Maloney, M., Giroux, N. (2001) Learning Rates with Direct Instruction, Precision Teaching and the Corrective Reading Series. *Journal of Precision Teaching & Celeration*. Vol.17 (2), pp. 89-91. Criteria 4
 18. Morgan, P L., Sideridis, G D (2006) Contrasting the Effectiveness of Fluency Interventions with or At Risk for Learning Disabilities: A Multilevel Random Coefficient Modeling Meta-Analysis. *Learning Disabilities Research & Practice*. Vol. 21 (4), pp. 191-210. Criteria 5
 19. Doughty, S. S., Chase, P. N., & Shields, E. M. O. (2004). Effects of Rate Building on Fluent Performance : A Review and Commentary, 1(1), 7–23. Criteria 5
 20. O'Donnell, P., Weber, K P., McLaughlin, T F. (2003) Improving correct and error rate and reading comprehension using key words and previewing: A Case report with a language minority report. *Education & Treatment of Children*. Vol.26 (3), pp. 237-254. Criteria 5
 21. White, L. (2002) Rick's reading progress. *Journal of Precision Teaching & Celeration*. Vol.18 (2), pp. 69-77. Criteria 5
 22. Nitti, J. (1990) Utilizing Precision Teaching to measure growth of reading comprehension skills in low achieving students. *Dissertation*. Criteria 1
 23. Binder, B. C. (1988) Precision Teaching : Measuring and attaining exemplary Precision Teaching : Measuring and Attaining Exemplary Academic Achievement, 10, 12–15. Criteria 4
 24. Sweeney, W. J., Ring M.M., Malanga, P. (2003) Using Curriculum-Based Assessment and Repeated Practice Instructional Procedures Combined with Daily Goal Setting to Improve Elementary Students Oral Reading Fluency: A Pre service Criteria 4 & 2

- Teacher Training Approach. *Journal of Precision Teaching & Celeration* Vol.19 (1), pp. 2-19.
25. Peterson, S.K., Scott, J., Sroka, K (1990) Using the language experience approach with precision. *Teaching Exceptional Children*. Spring 1990; 22, 3; Criteria 2
26. Binder, C ., & Watkins , C . L . (1990). Precision Teaching and Direct Instruction : Measurably superior instructional technology in schools . *Performance Improvement Quarterly* , 3 (4) , 74-96 ., 3. Criteria 4
27. Ardoin, S. P., Eckert, T. L., & Cole, . C. A. S. (2008). Promoting Generalization of Reading : A Comparison of Two Fluency-Based Interventions for Improving General Education Student ' s Oral Reading Rate, 237–252. Criteria 4
-

Appendix 2: Mapping the field

Roberts & Norwich (2010)				
Participants	Intervention	Design	Measures	Outcome
<p>77 students, across 5 Secondary schools in the South of England. 3 mainstream schools and 2 specialist. Chosen due to needing additional support with reading.</p> <p>Age range: From 11 years 4 months to 16 years and 2 months (mean: 13 years 6 months).</p>	<p>A 5-6 week Precision Teaching Intervention.</p> <p>Aim: To improve the word reading skills of students.</p> <p>Frequency: Daily sessions involving a procedure summarised as ‘teach–test–chart–review’ undertaken by TA’s</p>	<p>Quasi experimental design. Students randomly assigned into intervention or control group usual teaching arrangements (UTA). A wait list control group was used. In cohort 1, Group A received the intervention first, Group B received UTA and then the interventions were reversed. In cohort 2, Group C received the intervention first and Group D received UTA and then the interventions were reversed. Intervention was additional to UTA as it took place during break times, lunchtimes or afterschool.</p>	<p>Students were measured using the Reading Accuracy Measure (RAM), pre, mid and post interventions based on the 332 commonly occurring sight words from the National Literacy Strategy (NLS).</p> <p>The outcome of this assessment informed the PT programme plan and the same measure with used post intervention</p>	<p>Cohort 1 showed overall gains in their performance on the RAM at the three points of measurement, however Group A showed more of a gain when under UTA conditions.</p> <p>Cohort 2 showed gains in their performance of on the RAM and showed greater gains during periods with PT intervention took place.</p>

Hughes et al (2007)				
Participants	Intervention	Design	Measures	Outcome
<p>7 primary school aged participants whose reading age was significantly below their average peers.</p> <p>Age range: Intervention group 11 years 6 months to 12 years 8 months (Mean: 12 years 2 months). Age range for TAU group was 11 years 7 months to 11 years 11 months (Mean: 11 years 9 months).</p>	<p>10 week Precision Teaching intervention with 5 pupils. 2 participants received TAU.</p> <p>Aim: to improve the word reading frequency level of pupils</p> <p>Frequency: Intervention group received 20 minute session 3-4 times per week undertaken by a TA. TAU group received 20 minute sessions 3-4 times per week accessing the usual support from a TA.</p>	<p>Single Case Design. Pupils were randomly assigned to intervention and TAU group.</p>	<p>The two standardised tests were used: Group Reading Test II (GRT II) (NFER-Nelson, 2000) and the Vocabulary Scale of the Middle Years Information System (MidYIS) year 7 (CEM, 2007)</p> <p>Participants tested two months before the PT programme began and were retested one month after the PT programme</p>	<p>The intervention group showed improvement in their word reading over the 10-week period reading. The intervention group increased by a mean reading age of between 3-5 months (if participant 1 is excluded)</p> <p>The TAU group's mean reading age decreased by 1 month.</p> <p>Inter-observer reliability was calculated as 97.5% across all tests</p>

Lambe et al (2015)				
Participants	Intervention	Design	Measures	Outcome
7 typically developing 7-8 year olds. 4 girls and 3 boys with mean age of 7 years and 10months, ranging from 7 years 2monthsh to 8 years 4 months.	A 6 week Precision Teaching intervention. Intervention 20-30 minutes 3 days per week.	Concurrent Multiple Baseline Design (MBD). Pre-intervention–post-intervention fluency scores were recorded. Concurrent baseline data collections were commenced with the seven participants, and a PT intervention was introduced with each participant at staggered intervals according to whether the participants’ baseline data were stable. The aim was to demonstrate that change occurred contingent on the introduction of a PT intervention with each participant. Mean fluency data were also recorded pre-intervention and at follow-up.	MBD at the beginning of the intervention using Say All Fast a Minute Every Day Shuffled (SAFMEDS) fluency cards (Phase 1) and a Dolch story (Phase 2).	The MBD data show increased correct responding for All 7 participants showed increased word reading fluency from baseline to the end of each intervention phase. These improvements were maintained at a 3-week post-intervention follow-up. Mean inter observer agreement was 99.5% on words read correctly.

Downer (2007)				
Participants	Intervention	Design	Measures	Outcome
53 participants across 7 schools (1 infant, 4 junior, and 2 secondary) were selected by their teachers due to concern regarding their very slow progress relative to peers. There were more boys than girls.	<p>Precision Teaching intervention lasting between 2-26 weeks. The average period of exposure was approximately 22 weeks.</p> <p>Frequency: 4 minutes per day by trained TA's.</p> <p>The importance of the children being encouraged to practise their target words at home with parents was also emphasised.</p>	Single case (Multiple Baseline Design)	<p>Infant and junior schools: outcome measure was the increment in the number of words the child could instantly recognise.</p> <p>Secondary school children: outcome measure was the incremental change in reading age of the participants before and after exposure to the programme (The Salford Sentence Reading Test, 1991).</p>	<p>All the pupils made progress Junior boys particularly benefited from the intervention.</p> <p>Weekly increment in number of new words read across all ages ranged from 0.57 to 3.55 with an overall average of 2.15.</p> <p>Mean weekly increment's (words recognised) were: Year1= 2.30 Year 2= 1.97. Years 3-6= 3.61 Year 7= 3.91 Year 8= 3.96</p> <p>These mean incremental gains were all statistically highly significant (p<.001).</p>

Griffin & Murtagh (2015)				
Participants	Intervention	Design	Measures	Outcome
<p>40 participants (20 intervention group, 20 control) in Irish primary schools. Intervention group were chosen due to reading age being below their chronological ages. The Control group were pupils who were not receiving any specialist reading support but whose reading age was in line with their chronological age; therefore no group was deemed to be at a disadvantage. Groups were matched in terms of class-level and gender</p> <p>Age range: 7 years 4 months to 12 years 2 months [mean age nine years seven months].</p>	<p>An 8 week Precision Teaching intervention.</p> <p>Frequency: 10 minutes per day one to one with a Learning Support Teachers.</p>	<p>Mixed methods, quasi-experimental design and Qualitative interviews. The quantitative component involved a pre-post experimental design using an experimental and control group.</p> <p>Following intervention, a mixed between-within subjects' analysis of variance was carried out.</p>	<p>Both groups undertook the SNIPPT- Pack (Smart & Smart, 2007) pre and post. The experimental group (only) was also individually tested using 'Form A' of the GORT-4 (Wiederholt & Bryant, 2001) pre and post to establish reading accuracy, rate, fluency, comprehension and an overall reading ability composite score.</p> <p>Teachers and experimental group pupils were interviewed to obtain their views regarding the strengths and weaknesses of PT.</p>	<p>A differential growth in sight vocabulary scores across the control and intervention groups were found.</p> <p>Statistically significant large effect between Time 1 and Time 2 for the two groups (partial eta squared =.886).</p> <p>Large effect sizes were shown for the experimental group in the following areas: Reading accuracy (eta squared =.69) Reading fluency (eta squared=.64) Reading comprehension (eta squared =.55) Overall reading ability (eta squared =.72)</p> <p>Semi-structured interviews highlighted PT as an effective and motivational tool.</p>

Kessissoglou & Farrell (1995)				
Participants	Intervention	Design	Measures	Outcome
12 pupils, Years 5 and 6 in an inner city primary school in the UK. Participants identified as being at least 2 years behind their chronological age in reading.	<p>Precision Teaching intervention, implemented on a 1:1 basis with one of the authors (Kessissoglou). Intervention was over a 5 week period for 15 minutes each day. Reading accuracy and reading comprehension.</p> <p>Should be independent researcher/ interventionist</p>	<p>Participants were randomly assigned to one of three groups:</p> <p>Group 1: received PT intervention</p> <p>Group 2: spent 15 minutes with one of the authors (Kessissoglou) talking about general school issues and playing educational games.</p> <p>Group 3: received no additional help.</p>	<p>The reading scores of participants was assessed using the NARA pre, post and at a 5 week follow up.</p>	<p>Group 1 increased their reading accuracy by 7 months and reading comprehension by 18 months.</p> <p>Group 2 increased their reading accuracy by 5 months and reading comprehension by 8 months.</p> <p>Group 3 increased their reading accuracy by 5 months and reading comprehension by 0 months.</p>

Appendix 3: Weight of Evidence Criteria

A. Quality of Methodology

Measures

Weighting	Description
High	Needs at least 3 out of 4: <ul style="list-style-type: none"> • Reported reliability of .85 or above for all primary outcomes relevant to the review question or used well referenced measures, • Use at least 2 methods of assessment • Use at least 2 sources of assessment. • Reported validity for all measures used or use a well referenced, standardised or norm-referenced measure.
Medium	Needs 2: <ul style="list-style-type: none"> • Reported reliability of 0.85 for some primary outcomes or use well referenced measures. • Use at least 2 methods of assessment <i>or</i> used at least 2 sources of assessment. (The above criteria must be met for 75% of primary outcomes)
Low	Needs at least 1: <ul style="list-style-type: none"> • Reported reliability of .50 or above for some primary outcomes or used well referenced measures. • Use at least 2 methods of assessment or used at least 2 sources of assessment. (The above criteria must be met for 50% of primary outcomes)

EITHER A) Comparison Group

Weighting	Descriptions
High	Used an active comparison group (Needs 2): <ul style="list-style-type: none"> • Demonstrated group equivalence • Evidence of equivalent mortality and low attrition
Medium	Used a 'no intervention' group (Needs 2): <ul style="list-style-type: none"> • Demonstrated group equivalence • Evidence of equivalent mortality and low attrition.
Low	A comparison group <ul style="list-style-type: none"> • One of the following criteria must be met: a) group equivalence established c) equivalent mortality with low attrition.

OR B) For Single Case Designs

Weighting	Descriptions
High	The study must have made use of a multiple baseline design or included at least three attempts to demonstrate intervention effect (Horner et al., 2005). The study must have included generalisation and maintenance phase with at least three data points for each phase.
Medium	The study must detail three attempts to demonstrate intervention effect, as well as either generalisation or maintenance phases with at least three data points for the phase.

Low The study may not have demonstrated intervention effect three times. Generalisation or maintenance data may not be included, or may be less than three data points.

Analysis

Weighting	Descriptions
High	Conducted an appropriate analysis: (needs all 3) <ul style="list-style-type: none"> Controlled for family wise error (when appropriate) Provided enough information for all effect sizes to be calculated Had a sufficiently large sample size
Medium	Conducted an appropriate analysis (needs 2 of 3) <ul style="list-style-type: none"> Provided enough information for some effect sizes to be calculated Had a sufficiently large sample size Conducted pre and post measures
Low	Conducted an appropriate analysis (needs 1) <ul style="list-style-type: none"> Provided enough information for some effect sizes to be calculated Had a sufficiently large sample size

In determining overall WOE A, each weight was assigned a score of:

- 3 for '**High**' scores
- 2 for '**Medium**' scores
- 1 for '**Low**' scores

Therefore ratings from *Measures + Comparison Group/ Single Case Design + Analysis* ÷ **3** = **WOE A**

N.B. Scores in the range '2.6' to '3' were awarded a 'high' weighting, those in the range '1.5' to '2.5' achieved a 'medium' weighting and scores falling at '1.4' or below were given a 'low' weighting.

B. Relevance of Methodology

Either A) Comparison group

Weighting	Descriptions
High	To receive a high weighting for methodological relevance the study must have used a random assignment of participants to either an intervention control or alternative intervention group. They must have demonstrated group equivalences. Pre and post-test measure must have been obtained for all groups and outcome measures.
Medium	To receive a medium weighting for methodological relevance the study must have demonstrated group equivalences. A comparison group must have been used. Pre and post-test measure must have been obtained for all groups involved.
Low	To receive a low weighting, pre and post-test measure must have been obtained for all groups involved.

Or B) For Single Case Design

Weighting	Descriptions
High	In order to receive a high weighting for methodological relevance the study must have made use of a multiple baseline design or included at least three attempts to demonstrate intervention effect (Horner et al., 2005). The study must have included generalisation and maintenance phase with at least three data points for each phase.
Medium	In order to receive a medium weighting for methodological relevance the study must detail three attempts to demonstrate intervention effect, as well as either generalisation or maintenance phases with at least three data points for the phase.
Low	In order to receive a low weighting for methodological relevance the study may not have demonstrated intervention effect three times. Generalisation or maintenance data may not be included, or may be less than three data points.

C. Relevance of evidence to the review question

Weighting	Descriptions
High	Needs at least 4 of 5: <ul style="list-style-type: none"> • The intervention took place at least 4 times per week. • The programme ran for at least 10 weeks • The study reported a high level of fidelity to the intervention • The intervention was manualised. • Staff delivering the intervention were provided with training.
Medium	Needs at least 3 of 5: <ul style="list-style-type: none"> • The intervention took place 2 or 3 times per week • The programme ran between 4-10 weeks • The study reported the level of fidelity to the intervention • The study met at least 1 of the following conditions: <ul style="list-style-type: none"> • The intervention was manualised, or • Staff delivering the intervention were provided with training.
Low	Needs at least 2 of 3: <ul style="list-style-type: none"> • The intervention took place twice per week or less • The programme ran for at least 4 weeks • The fidelity to the programme was reported.

D. Overall weight of evidence

In determining overall methodological quality, each weight was assigned a score of:

- 3 for 'High' scores
- 2 for 'Medium' scores
- 1 for 'Low' scores

$$(WOE A + WOE B + WOE C) \div 3 = WOE D$$

N.B. Scores in the range '2.6' to '3' were awarded a 'high' weighting, those in the range '1.5' to '2.5' achieved a 'medium' weighting and scores falling at '1.4' or below were given a 'low' weighting.

Appendix 4: Coding Protocols

[adapted from Task Force on Evidence-Based Interventions in School Psychology, American Psychology Association, Kratochwill, T.R. (2003)]

Coding Protocol

Name of Coder: X

Date: 16.01.16

Full Study Reference in proper format:

Roberts, W., & Norwich, B. (2010). Using precision teaching to enhance the word reading skills and academic self-concept of secondary school students: a role for professional educational psychologists. *Educational Psychology in Practice*, 26(3), 279–298.
<http://doi.org/10.1080/02667363.2010.495215>

Intervention Name (description of study): Precision Teaching

Study ID Number: 1

Type of Publication:

Book/Monograph

Journal Article

Book Chapter

Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)

- Completely randomized design
- Randomized block design (between participants, e.g., matched classrooms)
- Randomized block design (within participants)
- Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)

- Nonrandomized design
- Nonrandomized block design (between participants)
- Nonrandomized block design (within participants)
- Nonrandomized hierarchical design

Optional coding for Quasi-experimental designs

A3. Overall confidence of judgment on how participants were assigned (select one of the following)

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)
- Very high (explicitly stated)
- N/A
- Unknown/unable to code

B Participants

Total size of sample (start of study): 77

Intervention group sample size: 77 (53 boys and 24 girls)

Control group sample size: Wait list control group. All participants had the intervention and TAU.

C. Type of Program

- Universal prevention program

Selective prevention program

Targeted prevention program

Intervention/Treatment

Unknown

D. Stage of Program

Model/demonstration programs

Early stage programs

Established/institutionalized programs

Unknown

E. Concurrent or Historical Intervention Exposure

Current exposure

Prior exposure

Unknown

Section 2 Key Features for Coding Studies and Rating Level of Evidence/Support

A Measurement (Estimating the quality of the measures used to establish effects)

(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes (see following table for a detailed breakdown on the outcomes)

Yes

No

Unknown/unable to code

A2 Multi-method (at least two assessment methods used)

Yes

No

N/A

Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

Yes

No

N/A

Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

Yes validated with specific target group

In part, validated for general population only

No

Unknown/unable to code

Overall Rating of Measurement: 3 2 1 0

B Comparison Group

B1 Type of Comparison group

Typical intervention

Attention placebo

Intervention element placebo

Alternative intervention

Pharmacotherapy

No intervention

Wait list/delayed intervention

Minimal contact

Unable to identify type of comparison

B2 Overall rating of judgment of type of comparison group

Very low

- Low
- Moderate
- High (strong inference)
- Very high
- Unable to identify comparison group

B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)

- By change agent
- Statistical (analyse includes a test for intervention)
- Other
- Not reported/None

B4 Group equivalence established

- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

B5 Equivalent mortality

- Low attrition (less than 20 % for post)
- Low attrition (less than 30% for follow-up)
- Intent to intervene analysis carried out?

Findings_____

Overall Level of Evidence ____

Overall Rating of Analysis: 3 2 1 0

C Appropriate Statistical Analysis

Analysis

1 _____
—

Appropriate unit of analysis

Familywise/experimenter wise error rate controlled when applicable

Sufficiently large N

Overall Rating of Analysis: 3 2 1 0

Quality Indicators: Within Single-Subject Research (Horner 2005)

Name of coder: X

Date: 16.01.16

Study number: 3

Full study:

Lambe, D., Murphy, C. & Kelly, M.E. (2015) The Impact of a Precision Teaching Intervention on the Reading Fluency of Typically Developing Children. *Behavioural Interventions*. 30: 364-377

Description of Participants and setting

	Yes	No
<ul style="list-style-type: none"> Participants are described with sufficient detail to allow others to select individuals with similar characteristic; (e.g., age, gender, disability, diagnosis). <i>Gender was not noted</i> 		No
<ul style="list-style-type: none"> The process for selecting participants is described with replicable precision. 		No
<ul style="list-style-type: none"> Critical features of the physical setting are described with sufficient precision to allow replication. 	Yes	

Dependent Variable

	Yes	No
<ul style="list-style-type: none"> Dependent variables are described with operational precision. 		No
<ul style="list-style-type: none"> Each dependent variable is measured with a procedure that generates a quantifiable index. 	Yes	
<ul style="list-style-type: none"> Measurement of the dependent variable is valid and described with replicable precision. 		No
<ul style="list-style-type: none"> Dependent variables are measured repeatedly over time. 	Yes	
<ul style="list-style-type: none"> Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet minimal standards {e.g., IOA = 80%; Kappa = 60%}. 	Yes	

Independent Variable

	Yes	No
<ul style="list-style-type: none"> Independent variable is described with replicable precision. 	Yes	
<ul style="list-style-type: none"> Independent variable is systematically manipulated and under the control of the experimenter. 	Yes	
<ul style="list-style-type: none"> Overt measurement of the fidelity of implementation for the independent variable is highly desirable 	Yes	

Baseline

	Yes	No
<ul style="list-style-type: none"> The majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur. 	Yes	
<ul style="list-style-type: none"> Baseline conditions are described with replicable precision. 	Yes	

Experimental Control/internal Validity

	Yes	No
The design provides at least three demonstrations of experimental effect at three different points in time.	Yes	
The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).		No
The results document a pattern that demonstrates experimental control.	Yes	

External Validity

	Yes	No
Experimental effects are replicated across participants, settings, or materials to establish external validity.	Yes	

Social Validity

	Yes	No
The dependent variable is socially important.	Yes	
The magnitude of change in the dependent variable resulting from the intervention is socially important.	Yes	
Implementation of the independent variable is practical and cost effective.	Yes	
Social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts.	Yes	