

Case Study 1: An Evidence-Based Practice Review Report

Theme: School/setting Based Interventions for Learning.

Is Reciprocal Teaching an effective intervention for developing the reading comprehension abilities of Secondary, Sixth Form and Further Education learners of English as an additional language (EAL) in the UK?

Summary

Reciprocal Teaching is an intervention which aims to improve the comprehension abilities of struggling readers (Palincsar & Brown, 1984). Using the three core principles of Proleptic Teaching (Brown, Campione, Reeve, Ferrara and Palincsar, 1991), Scaffolding (Wood, Bruner and Ross, 1976) and the Zone of Proximal Development (Vygotsky, 1978), Palincsar and Brown (1984) suggest four strategies for readers to learn: 1) summarising, 2) questioning, 3) clarifying and 4) predicting. Having a large body of research supporting its effectiveness (Alfassi, 2016; Rosenshine & Meister, 1994; Spörer, Brunstein and Kieschke, 2009), this systematic review aimed to investigate its effectiveness for learners of English as an Additional Language (EAL) in secondary, sixth form and further educational settings. A systematic literature search identified eight studies that met the inclusion criteria which were subsequently evaluated using the Weight of Evidence framework (Gough, 2007). Three out of the eight studies were given a rating of 'medium' and five were rated as 'low'. Despite the positive results of each study, this review concludes that it is not possible to argue for the effectiveness of Reciprocal Teaching for the EAL population in the UK based on the evidence available. Recommendations for future research are suggested.

Introduction**Reciprocal Teaching**

Research has recognised that one of the primary goals of the education system is to develop reading comprehension abilities (Spörer et al., 2009) as they are fundamental in being able to access knowledge. It has also been found that there is a correlation between reading ability and academic achievement and self-esteem (Hisken, 2011) which further highlights the importance of reading ability. Reciprocal Teaching (Palincsar & Brown, 1984) is one intervention that has been developed to address difficulties in reading comprehension ability. Palincsar and Brown (1984) recognise that reading comprehension, provided decoding abilities are adequate, is the product of three main factors: 1) considerate and appropriate texts, 2) the compatibility between the content and the reader's knowledge and 3) the strategies the reader actively applies during reading to enhance their understanding, retention and ability to avoid comprehension errors. Reciprocal Teaching aims to address the third factor they identify through the explicit teaching of metacognitive strategies.

Palincsar and Brown (1984) explain how competent readers flow from an automatic state of reading to a strategic state of reading. They explain that 'trigger' events during reading alert the competent reader to a comprehension error at which point the strategic state begins where the reader applies strategies to ameliorate the error. Once in the strategic state, the competent reader slows down significantly to allow extra processing to resolve the problem. However, for the less competent reader these 'triggers' may not be salient enough to alter the state of reading and even if they are salient enough the reader may not have the strategies in place to fully comprehend the text. Palincsar and Brown (1984) therefore identify four strategies to teach beginning readers: 1) summarising, 2) questioning, 3) clarifying and 4) predicting which they argue will give beginner readers the tools necessary to

improve their comprehension abilities. The four strategies should give the reader the ability to enhance and assess their comprehension.

The practical application of Reciprocal Teaching is built upon the three core principles of Proleptic Teaching (Brown et al., 1991), Scaffolding (Wood et al., 1976) and the Zone of Proximal Development (ZPD) (Vygotsky, 1978). The concept of Proleptic Teaching is the idea of teaching in expectation of competence and therefore having high expectations (Brown et al., 1991). The ZPD is defined as the area between where a learner can perform currently and where they can perform with expert help (Vygotsky, 1978) and it is theorised as the area where they will make most progress. The principles of Scaffolding are entwined with the ZPD where scaffolding is used to keep the learner in their ZPD. Scaffolds are supports that enable a learner to be successful through lowering the difficulty of the task to the point at which they are in their ZPD – the scaffolds are removed as a learner increases in competence. The role of the teacher in Reciprocal Teaching is to keep the learner in their ZPD through using Scaffolding and Proleptic Teaching (Palincsar & Brown, 1984). Reciprocal Teaching has been found to be effective when compared to alternative interventions (Spörer et al., 2009) and through a meta-analysis (Rosenshine & Meister, 1994) with effect sizes of (Cohen's) $d=0.32$ for standardised measures of comprehension and $d=0.88$ for researcher-developed measures of comprehension and for more specific populations such as those with learning difficulties (Alfassi, 2016; Gersten, Fuchs, Williams and Baker, 2001).

English as an Additional Language

English as an Additional Language (EAL) learners are defined as those who are school aged whose primary language is not English – in other parts of the world these students may be known by different labels such as English Language Learners (ELL) or English as a Second Language (ESL) learners (Schneider, 2014).

These individuals are not a homogenous group and may be at differing stages in their English development. According to the most recent (January 2019) School Census carried out in the UK, 21.2% of primary aged pupils and 16.9% of secondary aged pupils are classified as EAL learners which equates to almost 1.6 million students (*Statistics, 2019*).

Researchers have investigated the link between English proficiency and academic outcomes (GCSE results) for students aged 16 in an inner London setting and found that the students' stages of English fluency, which are linked with reading ability, were strongly associated with academic outcomes (Demie & Strand, 2006). The students at stages one (new to English) and two (becoming familiar with English) had a strong negative association, of around one standard deviation, to academic outcomes (Demie & Strand, 2006).

For 'late arriving students', those who enter the UK schooling system in Key Stage 4, which would be college or sixth form level, academic outcomes are extremely low (Hutchinson, 2018). The report also found that there are large differences between language groups with certain groups, such as Pashto speakers in Key Stage 4 (KS4), falling significantly further behind than others (Hutchinson, 2018). The report also highlights the lack of capacity of educational institutions, which also varies regionally, to support later arriving EAL students, particularly in KS4. Given the exceptional difficulties that later arriving EAL students face in general academic achievement, this systematic literature review will focus on secondary, sixth form and college age students.

In a three-year longitudinal study, it was found that EAL students had similar levels of reading accuracy to their non-EAL peers but were significantly lower on their scores of vocabulary and reading comprehension at every point throughout the study (Hutchinson et al., 2003). The authors of this study also suggest that students learning EAL typically have a two-year lag in their comprehension skills compared to

non-EAL learning students. This highlights the need for evidence-based interventions that are effective at aiding EAL learners to make further progress in their reading comprehension abilities. This need is backed by the report commissioned by The Bell Foundation (Hutchinson, 2018) that calls for further research into teaching and learning strategies that are effective for EAL learners (Schneider, 2014) – which leads to the review question: *Is Reciprocal Teaching an effective intervention for developing the reading comprehension abilities of Secondary, Sixth Form and Further Education learners of English as an additional language (EAL) in the UK?*

Critical Review of Evidence

Systematic Literature Search

A systematic literature search was conducted on the 8th of December 2019 which produced 87 results. The search terms can be seen in Table 1 and the results for the individual databases can be seen in Table 2. After duplicates were removed (n=41), the articles were screened at title and abstract level using the inclusion and exclusion criteria, which can be seen in Table 3. This resulted in 10 articles being screened at full text level and 8 being selected for inclusion in the systematic review. The two studies excluded at full text screening along with exclusion reason can be seen in Table 4. The full literature search process can be visualised using the flow diagram in Figure 1. Full references for the 36 articles excluded at title and abstract screening can be seen in appendix 5.

Table 1

Search terms

("reciprocal reading" OR "reciprocal teaching") AND ("english as an additional language" OR "english as a foreign language" OR "english as a second language" OR "english language learners" OR "EAL" OR "EFL" OR "ESL" OR "ELL")

Table 2

Search results

ERIC (EBSCO)	British Education Index (EBSCO)	Teacher Reference Centre (EBSCO)	PsycInfo	ERIC (ProQuest)	Scopus	Total
39	5	4	8	20	11	87

Table 3

Inclusion and exclusion criteria with rationale

	Inclusion criteria	Exclusion criteria	Rationale
1. Type of articles	Published in a peer reviewed journal	Not in a peer reviewed journal	To ensure the quality of the study as it has been checked by peers in the research community
2. Type of data	Primary quantitative data	Not primary quantitative data	Need quantitative data in order to objectively compare the studies
3. Intervention	Reciprocal Teaching/Reading principles are used	No Reciprocal Teaching/reading principles are used	This is the topic of interest
4. Participant	Participants are defined as learning	Participants are not defined as learning	This is the group of interest to the study

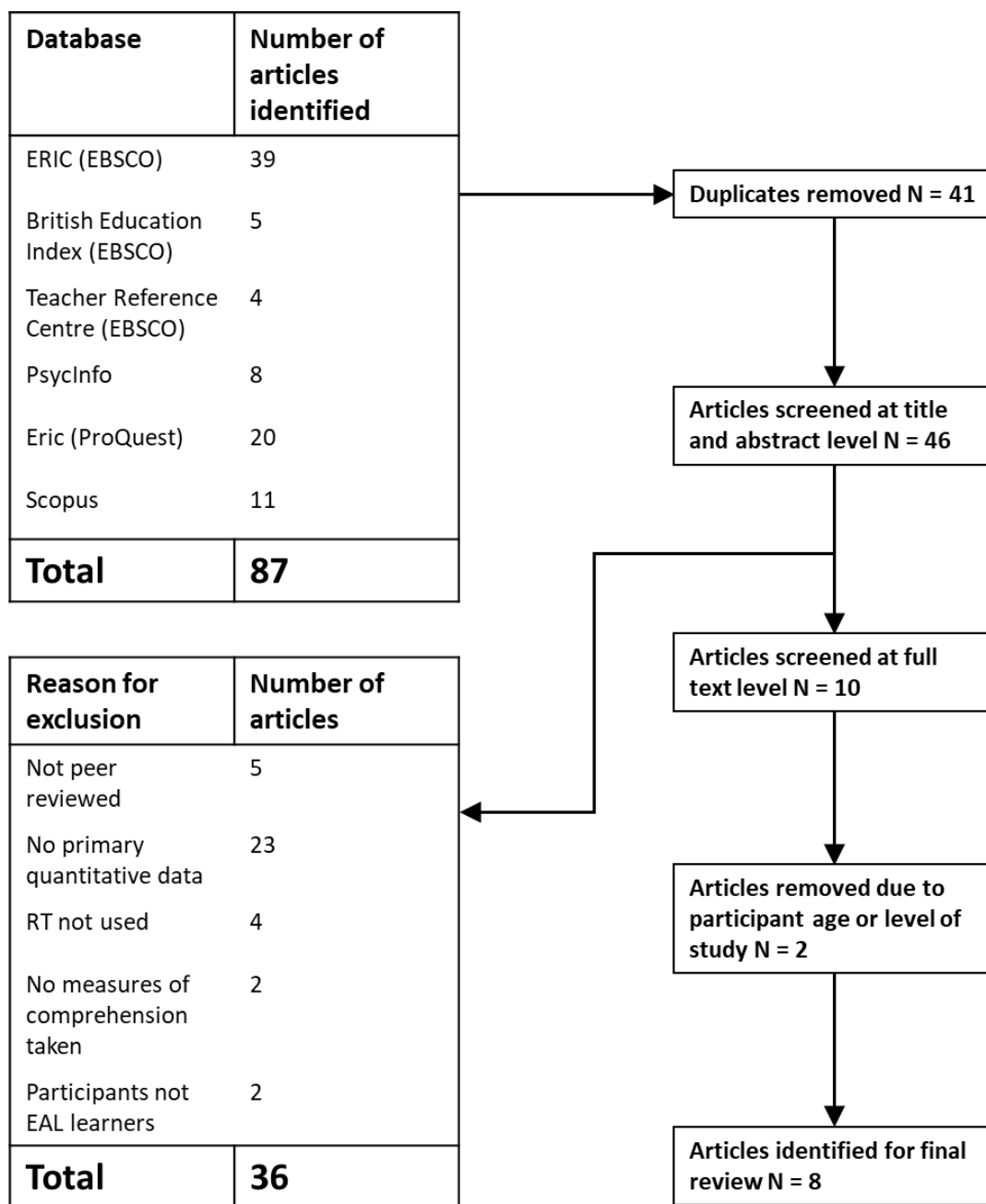
	Inclusion criteria	Exclusion criteria	Rationale
language status	English as either a second or additional language	English as a second or additional language	as there has been no systematic review of the literature surrounding the effectiveness of Reciprocal Teaching for EAL learners
5. Measures	A measure of comprehension is taken pre and post intervention	No measure of comprehension taken pre and post intervention	This is the dependent variable and will allow for comparison before and after intervention.
6. Participant age	The sample is within the 11-19 age group OR are reported as being in the first year of undergraduate studies.	The participants are younger than 11, older than 19 OR are reported as being in second year of undergraduate studies or later	The review question pertains to secondary, sixth-form and further education which is from 11 to 19 years of age. First year of undergraduate studies typically is up to the age of 19

Table 4

Studies excluded at full text screening with exclusion criteria

Study	Exclusion criteria
Izadi, M., & Nowrouzi, H. (2016). Reciprocal Teaching and Emotional Intelligence: A Study of Iranian EFL Learners' Reading Comprehension. <i>Reading Matrix: An International Online Journal</i> , 16(1), 133–147.	Participant age
Navaie, L. A. (2018). The Effects of Reciprocal Teaching on Reading Comprehension of Iranian EFL Learners. <i>Advances in Language and Literary Studies</i> , 9(4), 26–30.	Participant age

Figure 1 - Flow diagram of the literature search and screening process



Mapping the Field

Full references of the 8 included studies can be seen in Table 5.

Table 5

List of included studies

-
- Choo, T. O. L., Eng, T. K., & Ahmad, N. (2011). Effects of Reciprocal Teaching Strategies on Reading Comprehension. *Reading Matrix: An International Online Journal*, 11(2), 140–149.
- Dabarera, C., Renandya, W. A., & Zhang, L. J. (2014). The impact of metacognitive scaffolding and monitoring on reading comprehension. *System*, 42, 462–473.
- Fung, I. Y. Y., Wilkinson, I. A. G., & Moore, D. W. (2002). L-1-Assisted Reciprocal Teaching To Improve ESL Students' Comprehension of English Expository Text. *Learning and Instruction*, 13(1), 1–31.
- Huang, C.-T., & Yang, S. C. (2015). Effects of Online Reciprocal Teaching on Reading Strategies, Comprehension, Self-Efficacy, and Motivation. *Journal of Educational Computing Research*, 52(3), 381–407.
- Klingner, J., & Vaughn, S. (1996). Reciprocal Teaching of Reading Comprehension Strategies for Students with Learning Disabilities Who Use English as a Second Language. *The Elementary School Journal*, 96(3), 275-293.
- Soonthornmanee, R. (2002). The Effect of the Reciprocal Teaching Approach On the Reading Comprehension of Efl Students. *RELC Journal*, 33(2), 125–141.
- Tseng, S., & Yeh, H. (2018). Integrating reciprocal teaching in an online environment with an annotation feature to enhance low-achieving students' English reading comprehension. *Interactive Learning Environments*, 26(6), 789-802.
- Yeh, H.-C., Hung, H.-T., & Chiang, Y.-H. (2017). The Use of Online Annotations in Reading Instruction and Its Impact on Students' Reading Progress and Processes. *ReCALL*, 29(1), 22–38.
-

In Table 6 key details of the eight included studies have been summarised.

Table 6

Mapping the field

Study	Country	Participants	Design	How was Reciprocal Teaching used?	Measures
Choo, Eng and Ahmad (2011)	Malaysia	68 'low-proficiency' sixth form students	Quasi-experimental – 2 classes randomly assigned to either experimental or control conditions	9 lessons over one month	Researcher devised measure of reading comprehension based on Malaysian University Entrance Exam
Dabarera, Renandya and Zhang (2014)	Singapore	67 'express-stream' students in Secondary 1 in a government run school. Mean age 12.1 years	Quasi-experimental with two classes randomly assigned to experimental or control conditions	5-week intervention with two 60-minute lessons each week (600 minutes in total)	University of New South Wales, International Competitions and Assessments for School English Comprehension Paper
Fung, Wilkinson and Moore (2002)	New Zealand	12 students across years 7 and 8. Age range from 11.6 to 13.6. Participants could read at age expected levels in Mandarin and were between 4 and 6 years behind in English reading ability	Single-subject multiple baseline design	35 minutes each day for 20 days in school one (700 minutes). 35 minutes each day for 15 days in school two (525 minutes). Lessons alternated between Mandarin and English	Neal Analysis of Reading Ability

Study	Country	Participants	Design	How was Reciprocal Teaching used?	Measures
Huang and Yang (2015)	Taiwan	36 remedial English learners in their first week of university	Quasi-experimental with the bottom 36 of 108 students randomly allocated to experimental and control groups	Reciprocal Teaching principles were carried out online via video conferencing software in 10 ninety-minute sessions (900 minutes)	Reading comprehension measure designed by the researchers but based on a Taiwanese General English Proficiency Test
Klinger and Vaughn (1996)	United States of America	26 7 th and 8 th grade students with IQ scores 1.5 standard deviations or more below average. Participants were all defined as having a learning disability	One group pre-test post-test design	40 minutes a day for 27 days (1080 minutes)	Gates-MacGinitie Reading Comprehension Test
Soonthornmanee (2002)	Thailand	84 first year university students	Quasi-experimental with 42 in experiment and 42 in control group	Equal amounts for each group across a single semester	Researcher adapted version of the Test of English as a Foreign Language (TOEFL)
Tseng and Yeh (2018)	Taiwan	22 low-achieving university students	One group pre-test post-test design	Online use of Reciprocal Teaching principles with small group work. 12-week intervention	Test of English for International Communication (TOIEC)

Study	Country	Participants	Design	How was Reciprocal Teaching used?	Measures
Yeh, Hung and Chiang (2017)	Taiwan	54 first year university students rated as 'low-intermediate' in their English ability	One group pre-test post-test design	Online use of Reciprocal Teaching Principles with group discussions. Weekly 2-hour sessions over 18 weeks (2160 minutes)	Test of English for International Communication (TOIEC)

Weight of Evidence

The eight included studies were evaluated using the Weight of Evidence framework (Gough, 2007) which is a tool for systematically evaluating the extent to which a set of evidence answers a given question. The framework includes three Weight of Evidence (WoE) areas labelled A, B and C which are applied to each study individually. WoE A assesses methodological quality, WoE B assesses methodological relevance and WoE C assesses the relevance of the evidence to the review question. The results are then combined into an overall measure of how useful that study is in answering the review question, known as WoE D.

In order to assess WoE A, each study, as they were all group intervention studies, was assessed using the quality indicators for group experiment and quasi-experiment research in special education (Gersten et al., 2005). This lists 10 essential criteria and 8 desirable criteria which studies need to meet to be considered acceptable or high quality. Table 7 lists the WoE A rating criteria. Table 8 provides a summary of the results of WoE A. Appendix 1 contains an example completed coding protocol for WoE A.

Table 7

Weight of Evidence A (WoE A) Criteria for Group Experimental and Quasi-Experimental Studies

WoE A Rating	Criteria
3 – High	9 essential criteria met and 4 or more desirable criteria met.
2 – Medium	9 essential criteria met and 1, 2 or 3 desirable criteria met.
1 – Low	Less than 9 essential criteria met.

Note – Weight of Evidence A criteria is based on guidance by Gersten et al. (2005)

Table 8

Summary of Weight of Evidence A (WoE A) Ratings for Included Studies

Study	Number of essential criteria met	Number of desirable criteria met	WoE A rating
Choo et al. (2011)	6	2	1 (low)
Dabarera et al. (2014)	9	2	2 (medium)
Fung et al. (2002)	6	3	1 (low)
Huang and Yang (2015)	8	2	1 (low)
Klinger and Vaughn (1996)	6	3	1 (low)
Soonthornmanee (2002)	5	3	1 (low)
Tseng and Yeh (2018)	6	3	1 (low)
Yeh et al. (2017)	5	1	1 (low)

Each study was then evaluated for WoE B as informed by the typologies of evidence (Petticrew & Roberts, 2003). The criteria for WoE B can be viewed in Appendix 2. Next, WoE C was assessed using criteria designed to assess relevance of the study to the review question. The criteria for WoE C can be seen in Appendix 3 and the individual scores for each study for each criterion can be seen in Appendix 4. The results were combined to create an overall WoE D which can be seen in Table 9 which provides a summary of all WoE ratings.

Table 9

Summary of Weight of Evidence Ratings (WoE) A, B, C and D.

Study	WoE A – Quality of Methodology	WoE B – Relevance of Methodology	WoE C – Relevance of Evidence to the Review Question	WoE D – Overall WoE Rating
Choo et al. (2011)	1 (Low)	2	2.2	5.2 (Medium)
Dabarera et al. (2014)	2 (Medium)	2	2.3	6.3 (Medium)
Fung et al. (2002)	1 (Low)	1	2.5	4.5 (Low)
Huang and Yang (2015)	1 (Low)	2	2	5 (Medium)
Klinger and Vaughn (1996)	1 (Low)	1	2.7	4.7 (Low)
Soonthornmanee (2002)	1 (Low)	2	1.7	4.7 (Low)
Tseng and Yeh (2018)	1 (Low)	1	2.2	4.2 (Low)
Yeh et al. (2017)	1 (Low)	1	2.2	4.2 (Low)
Average	1.13 (Low)	1.5	2.25	4.88 (Low)

Note – final WoE D calculated by summing across WoEs A, B and C. Low, medium and high ratings were decided upon based on the following ranges low (0 – 5), medium (5 – 7) and high (7 – 9).

Participants

A total of 369 participants from six countries were included in this review across the eight studies. The ages ranged from the beginning of secondary school with a lower range of 11.6 years old (Fung et al., 2002) to the start of university (Huang & Yang, 2015; Soonthornmanee, 2002; Yeh et al., 2017). Given the large age range across the studies, it is difficult to consider the findings highly relevant for any particular age group.

The participants were drawn from a range of international settings including secondary schools (Dabarera et al., 2014; Fung et al., 2002; Klinger & Vaughn, 1996), sixth form (Choo et al., 2011) and universities (Huang & Yang, 2015; Soonthornmanee, 2002; Tseng & Yeh, 2018; Yeh et al., 2017). This lack of representation from UK educational establishments was reflected in the WoE C ratings with studies set in foreign sixth forms and secondary settings being rated higher than those in university settings.

The language setting within which the participants were living was also considered with studies set in countries where the first language was English (Fung et al., 2002; Klinger & Vaughn, 1996) being rated higher. WoE C ratings also reflected participant level of English where all studies but one study (Choo et al., 2011) provided a quantifiable measure of participant level of English thus increasing the generalisability of findings overall.

Study Design

Four studies included control groups in their design along with evidence of the similarity of the control group to the experimental group (Choo et al., 2011; Dabarera et al., 2014; Huang & Yang, 2015; Soonthornmanee, 2002). There was however a large amount of variety in what the control groups did. In one study (Choo et al., 2011) it was reported that the control group studied the same texts only – suggesting no input from a teacher or person implementing an intervention.

Whereas the other studies utilised different methods for teaching the control group: Drill and Practice (Dabarera et al., 2014), Direct Instruction (Huang & Yang, 2015) and Skill-Based Method (Soonthornmanee, 2002). The study conducted by Soonthornmanee (2002) was particularly limited in its description of the method of teaching applied to the control group. The inclusion of a control group was reflected in higher WoE B ratings for these studies. The remaining four studies (Fung et al.,

2002; Klinger & Vaughn, 1996; Tseng & Yeh, 2018; Yeh et al., 2017) did not use control groups which was reflected in lower WoE B ratings. The Fung et al. (2002) study was described by its authors as a single subject design, despite this the Gersten et al. (2005) coding protocol for group experimental and quasi-experimental research was used. This was because of the large number of participants (n=12) for this type of design and the way in which the measures were taken and data presented allowed for the data to be treated as one group. The intention of the Fung et al. (2002) study was to assess the effectiveness of an intervention and as the sample was large enough to be treated as a group, allowing for it to be studied as if it was a pre-post single group design, it was deemed appropriate to use the Gersten et al. (2005) coding protocol for group research. None of the eight included studies used any follow up measures to assess longer term impact of the intervention which was reflected in lower WoE A ratings for all studies.

Intervention Implementation

In order for the results of this review to be highly applicable to UK educational settings, the ideal intervention study would be carried out using English in a country where English is the first language – only one study met this criteria (Klinger & Vaughn, 1996). Another study partially met this criteria as it was set in New Zealand, a country where English is the first language, but the intervention was carried out alternating between English and Mandarin (Fung et al., 2002). In terms of effect sizes, the study which alternated between the first language of the setting and the first language of the students, had a greater effect (see Table 10 in the 'Outcomes and Meta Analysis' section). The remaining six studies were carried out in countries where English is not the first language as reflected in WoE C ratings.

It is important to consider the person implementing the intervention in these studies as their proficiency in English and the student's first language affects the

generalisability of the findings to the UK education system. In each study the intervention was carried out by the researcher and in two cases the researcher was also self-defined as a teacher (Dabarera et al., 2014; Huang & Yang, 2015). No measures of the language ability of the person implementing the intervention were provided in any study which reduces the generalisability of the results. It is possible that having a highly competent bilingual teacher for EAL learners is the reason for the positive results found across the studies, however this cannot be factored in without sufficient information regarding the person implementing the intervention's language ability.

Reciprocal Teaching is designed to be carried out in small groups with a teacher informed about Reciprocal Teaching principles (Palincsar & Brown, 1984). Three studies deviated from this and investigated the effects of implementing Reciprocal Teaching principles through online interfaces (Huang & Yang, 2015; Tseng & Yeh, 2018; Yeh et al., 2017). The study conducted by Huang and Yang (2015) required participants to work in a university Multimedia Meeting Centre using a webcam and microphone to video conference with the instructor and peers. The studies by Tseng and Yeh (2018) and Yeh et al. (2017) required participants to use Google Docs to work collaboratively to annotate texts based on the Reciprocal Teaching principles and to discuss their annotations using an online chat room. These three studies which utilised an online approach still applied the Reciprocal Teaching principles combined with group collaboration but without the face to face interaction, so these studies were given lower ratings in WoE C.

The eight studies were characterised by a large range in the amount of time the intervention was carried out for – see the mapping table for details (Table 6). At the lower end of the range was the study by Dabarera et al. (2014) with 600 minutes of intervention time – this study also had the lowest effect size of the eight studies. The study with the greatest amount of intervention time included 2160 minutes of

intervention (Yeh et al., 2017) with the largest effect size of the studies where minutes of intervention could be calculated. There were studies where only basic details were provided about length of intervention. Soonthornmanee (2002) reported only that there was 'equal amounts for each group across a single semester' which makes it difficult to generalise from the findings. Similarly, Choo et al. (2011) reported that nine lessons took place over a month but not the length of the lessons which again reduces the generalisability. This is reflected in the WoE A ratings with those studies failing to report specifics receiving lower ratings. The large range of intervention time across the eight studies increases the incomparability of the results.

Measures

The eight included studies were characterised by a large range of measures of reading comprehension with lack of standardised measures of reading being utilised. Four studies (Fung et al., 2002; Klinger & Vaughn, 1996; Tseng & Yeh, 2018; Yeh et al., 2017) used standardised measures of reading ability (Neale Analysis of Reading Ability, Gates-MacGinitie Reading Comprehension Test and the Test of English for International Communication). Even within the four studies that did use standardised measures there was a wide range of tests used which reduces generalisability. The remaining four studies (Choo et al., 2011; Dabarera et al., 2014; Huang & Yang, 2015; Soonthornmanee, 2002) again used a variety of tests which were either researcher adapted (n=3) or non-standardised (n=1). See the mapping table (Table 6) for a summary of measures used. The range of measures used across the studies combined with the lack of detail reported where measures were researcher adapted mean that the findings cannot be considered highly reliable, this was reflected in WoE A and WoE C.

Additionally, it has been found that when standardised measures of reading comprehension are used, in a review of 16 Reciprocal Teaching studies not using an EAL sample, the median effect size was 0.32 (Rosenshine & Meister, 1994).

However, when researcher-developed measures of comprehension are used the median effect size was found to be 0.88 (Rosenshine & Meister, 1994).

Conceivably, this is because researchers are designing their tests around the Reciprocal Teaching principles which would therefore reduce the generalisability of their findings. This has been reflected in the WoE C ratings where using a researcher-developed measure of comprehension has been given a lower rating.

This resulted in four studies (Choo et al., 2011; Dabarera et al., 2014; Huang & Yang, 2015; Soonthornmanee, 2002) receiving lower ratings.

Outcomes and Meta Analysis

All studies reported statistically significant effects. A meta analysis of results was carried out in order to assess the overall effectiveness. Effect sizes and confidence intervals from each study were calculated using the Campbell Calculator (Wilson, n.d.) and inputted into Meta Essentials (Suurmond, Van Rhee and Hak, 2017) in order to calculate the overall effect size using a Random Effects Model. This model does not assume a 'true' effect size, rather it takes into account the fact that the included effects are a random sample of all possible effects which creates a more balanced way of calculating the overall summary effect by including between-studies variance which leads to a more cautious interpretation as reflected in a greater range between upper and lower confidence intervals (Borenstein, Hedges, Higgins and Rothstein, 2010). The eight included studies were split into two separate meta analyses due to the way that statistics were reported or study design. Meta analysis 1 (figure 2) included those studies where a control group could be factored into the effect size calculation (Dabarera et al., 2014; Huang & Yang, 2015;

Soonthornmanee, 2002) and meta analysis 2 (figure 3) included the remaining five studies where control groups could not be factored into the effect size calculations due to either insufficient data reporting or lack of control group. This split enables comparison across designs to see the impact of control groups. See Table 10 for details of all effect sizes, how effect sizes were calculated, confidence intervals and results of both meta analyses in numerical form. Also included in Table 10 below are the results of the combined meta analysis which included all studies.

Table 10

Results of meta analyses and all effect sizes with confidence intervals and accompanying WoE D ratings for reference.

Study	Sample size	Outcome measure	Effect Size (Cohen's d)	95% confidence interval lower	95% confidence interval upper	WoE D Rating
Choo et al. (2011)	68	Researcher devised measure of reading comprehension based on Malaysian University Entrance Exam	2.4057	1.7817	3.0298	5.2 (Medium)
Dabarera et al. (2014)	67	University of New South Wales, International Competitions and Assessments for School English Comprehension Paper	0.4191	-0.0762	0.9144	6.3 (Medium)
Fung et al. (2002)	12	Neal Analysis of Reading Ability	1.2864	0.0433	2.5295	4.5 (Low)
Huang and Yang (2015)	36	Reading comprehension measure designed by the researchers but based on a Taiwanese General English Proficiency Test	0.9752	0.2842	1.6663	5 (Medium)

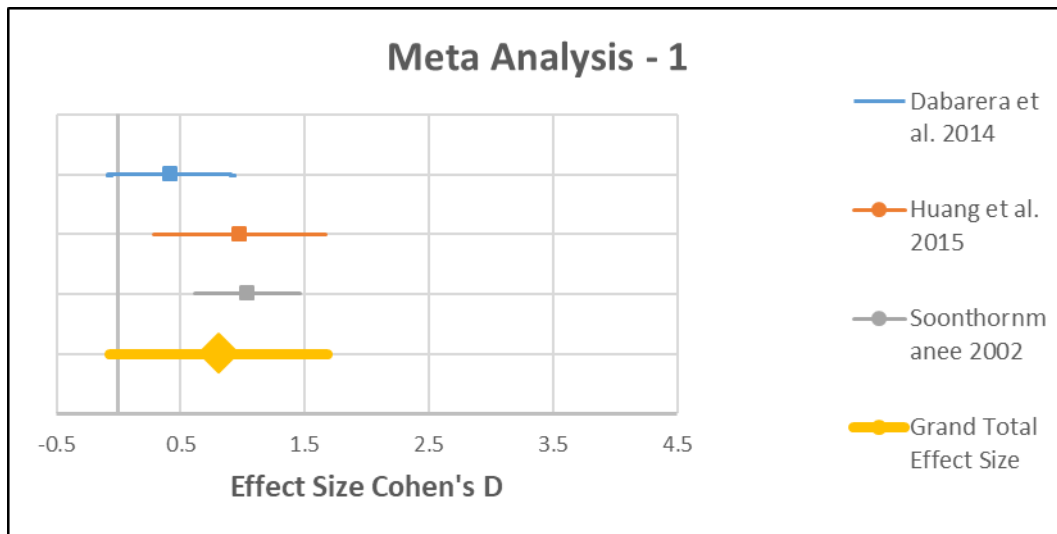
Study	Sample size	Outcome measure	Effect Size (Cohen's d)	95% confidence interval lower	95% confidence interval upper	WoE D Rating
Klinger and Vaughn (1996)	26	Gates-MacGinitie Reading Comprehension Test	0.5808	-0.204	1.3656	4.7 (Low)
Soonthornmanee (2002)	84	Researcher adapted version of the Test of English as a Foreign Language (TOEFL)	1.0412	0.6172	1.4652	4.7 (Low)
Tseng and Yeh (2018)	22	Test of English for International Communication (TOIEC)	2.7673	1.5981	3.9366	4.2 (Low)
Yeh et al. (2017)	54	Test of English for International Communication (TOIEC)	0.7911	0.3995	1.1828	4.2 (Low)
Meta analysis 1 Grand Total Effect Size	187		0.81	-0.07	1.68	5.3 (Medium)
Meta analysis 2 Grand Total Effect Size	182		1.52	0.32	2.73	4.6 (Low)
Combined meta analysis	369		1.2	0.52	1.89	4.88 (Low)

Note: for studies marked in bold effect sizes were calculated factoring in control groups.

Effect sizes for meta analysis 1 were calculated by dividing the mean pre-post difference in experimental and control groups by the pooled standard deviation of the pre-test scores.

Effect sizes for meta analysis 2 were calculated by dividing the pre-post mean difference by the standard deviation of the pre-test score.

Figure 2 - Meta Analysis 1 including only studies where control groups could be factored into effect size calculations



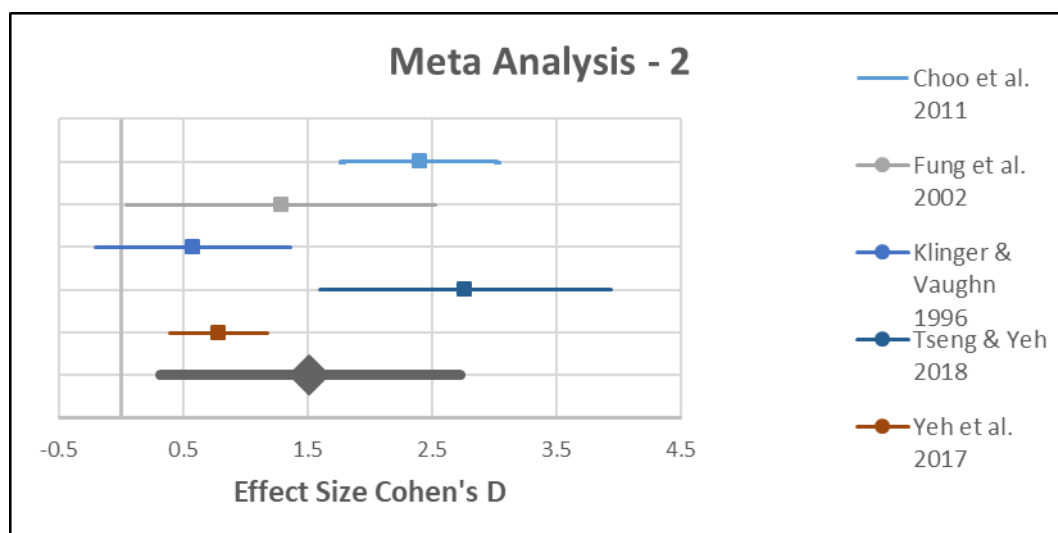
From meta analysis 1 (figure 2), where only those studies that allowed for control groups to be factored into effect size calculations, the Grand Total Effect Size is $d=0.81$, however the 95% confidence intervals span 0. Whilst $d=0.81$ may be considered large, see Table 11 (Cohen, 1992), with the confidence intervals spanning 0 it is difficult to conclusively assert that these three studies produced a positive effect.

Table 11

Cohen's descriptors of effect sizes

Small	Medium	Large
0.2	0.5	0.8

Figure 3 - Meta Analysis 2 including studies where control groups could not be factored into effect size calculations



Meta analysis 2 (figure 3), which includes the studies where control groups could not be factored into effect size calculations, presents a differing picture. The Grand Total Effect Size is $d=1.52$, which can be judged to be large, and the 95% confidence intervals do not span zero. This suggests that these studies have been effective in improving comprehension abilities for EAL students, however these results needed to be viewed with caution.

Without a control group it is not possible to rule out other potential factors that may have affected performance. As Reciprocal Teaching is carried out over time, without a control group it is impossible to rule out improvement over time or improvement due to attention effects for example.

When comparing meta analysis 1 to meta analysis 2 it is clear that control groups have a strong effect, for these specific studies, in reducing effect sizes which may be an argument for the 'true' effect size being closer to the $d=0.81$ of meta analysis 1 rather than the $d=1.52$ of meta analysis 2.

As it has been found that researcher developed measures of reading comprehension produce smaller effect sizes (Rosenshine & Meister, 1994), a third meta analysis was considered in order to compare those studies which used researcher developed measures of reading comprehension to those that did not. However, this would have left only two studies (Fung et al., 2002; Klinger & Vaughn, 1996) and neither of these studies had control groups so it was not deemed to be appropriate to carry it out.

Conclusion and Recommendations

Conclusion

This systematic review aimed to assess the effectiveness of the Reciprocal Teaching intervention for developing reading comprehension abilities for EAL learners in secondary, sixth form and further education settings. Eight studies met the inclusion criteria and were included for review. All eight studies showed the effectiveness of Reciprocal Teaching in developing reading comprehension abilities for EAL learners with meta analysis results of $d=0.81$ (for those where control groups could be factored in) and $d=1.52$ (for those where control groups could not be factored in). The large effects must be interpreted with caution as only three studies achieved a WoE D rating of 'medium' (Choo et al., 2011; Dabarera et al., 2014; Huang & Yang, 2015) with the remaining five studies achieving a WoE D rating of 'low'. Whilst the body of research supporting Reciprocal Teaching in more general contexts is compelling, based on these findings it is not possible to conclusively support Reciprocal Teaching for EAL learners in the UK. The evidence generally lacks methodological quality, lacks UK based studies, has a lack of clarity of the person implementing the interventions' language abilities and typically does not utilise standardised measures of reading comprehension.

Recommendations for Future Research

In order to answer the present research question fully, future research needs to be UK based for it to be highly generalisable to UK educational settings. Further to this, the language ability of the person implementing the intervention, in both their own language and in the participant's language, should be reported in detail as it is unclear what effect this has. For UK educational settings this would be valuable information when making staffing decisions about who will implement interventions.

Additionally, when considering research design future researchers need to ensure that control groups are included, as control groups are more effective at answering effectiveness questions (Petticrew & Roberts, 2003), and that information regarding control group treatment is detailed. Statistical reporting should also make it possible to calculate effect sizes that factor in control groups where control groups have been implemented.

Finally, future research in this area should consider the measures used to assess reading comprehension. Given that researcher developed measures have been found to increase effect sizes when studying Reciprocal Teaching (Rosenshine & Meister, 1994), one would expect that the improved validity and reliability that standardised tests provide would mean that all future research into Reciprocal Teaching would use standardised measures of reading, however this review has found that this isn't the case. A rejection of researcher developed or adapted measures of reading comprehension in future research into the effectiveness of Reciprocal Teaching for EAL learners is therefore advocated for.

References

- Alfassi, M. (2016). Reading for Meaning: The Efficacy of Reciprocal Teaching in Fostering Reading Comprehension in High School Students in Remedial Reading Classes: *American Educational Research Journal*.
<https://doi.org/10.3102/00028312035002309>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97–111. <https://doi.org/10.1002/jrsm.12>
- Brown, A. L., Campione, J. C., Reeve, R. A., Ferrara, R. A., & Palincsar, A. S. (1991). Interactive learning and individual understanding: The case of reading and mathematics. In *Culture, schooling, and psychological development* (pp. 136–170). Ablex Publishing.
- Choo, T. O. L., Eng, T. K., & Ahmad, N. (2011). Effects of Reciprocal Teaching Strategies on Reading Comprehension. *Reading Matrix: An International Online Journal, 11*(2), 140–149. eric.
- Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science, 1*(3), 98–101. <https://doi.org/10.1111/1467-8721.ep10768783>
- Dabarera, C., Renandya, W. A., & Zhang, L. J. (2014). The impact of metacognitive scaffolding and monitoring on reading comprehension. *System, 42*, 462–473. <https://doi.org/10.1016/j.system.2013.12.020>
- Demie, F., & Strand, S. (2006). English language acquisition and educational attainment at the end of secondary school. *Educational Studies, 32*(2), 215–231. <https://doi.org/10.1080/03055690600631119>
- Fung, I. Y. Y., Wilkinson, I. A. G., & Moore, D. W. (2002). L-1-Assisted Reciprocal Teaching To Improve ESL Students' Comprehension of English Expository Text. *Learning and Instruction, 13*(1), 1–31. ERIC.
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children, 71*(2), 149–164. <https://doi.org/10.1177/001440290507100202>
- Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching Reading Comprehension Strategies to Students With Learning Disabilities: A Review of Research. *Review of Educational Research, 71*(2), 279–320. <https://doi.org/10.3102/00346543071002279>
- 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. <https://doi.org/10.1080/02671520701296189>

- Hisken, L. J. (2011). The Correlation Between Self-Esteem and Student Reading Ability, Reading Level, and Academic Achievement. *University of Central Missouri*, 43.
- Huang, C.-T., & Yang, S. C. (2015). Effects of Online Reciprocal Teaching on Reading Strategies, Comprehension, Self-Efficacy, and Motivation. *Journal of Educational Computing Research*, 52(3), 381–407. <https://doi.org/10.1177/0735633115571924>
- Hutchinson, J. (2018). Educational Outcomes of Children with English as an Additional Language. *The Bell Foundation*, 43.
- Hutchinson, J. M., Whiteley, H. E., Smith, C. D., & Connors, L. (2003). The developmental progression of comprehension-related skills in children learning EAL. *Journal of Research in Reading*, 26(1), 19–32. <https://doi.org/10.1111/1467-9817.261003>
- Klinger, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities... *Elementary School Journal*, 96(3), 276. trh.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. *Cognition and Instruction*, 1(2), 117–175. https://doi.org/10.1207/s1532690xci0102_1
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology & Community Health*, 57(7), 527–529. <https://doi.org/10.1136/jech.57.7.527>
- Rosenshine, B., and Meister, C. (1994). Reciprocal Teaching: A Review of the Research. *Review of Educational Research*, 64(4), 479–530. <https://doi.org/10.3102/00346543064004479>
- Schneider, C. (2014). *School approaches to the education of EAL students: Language development, social integration and achievement*. The Bell Foundation.
- Soonthornmanee, R. (2002). The Effect of the Reciprocal Teaching Approach On the Reading Comprehension of Efl Students. *RELC Journal*, 33(2), 125–141. <https://doi.org/10.1177/003368820203300206>
- Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction*, 19(3), 272–286. <https://doi.org/10.1016/j.learninstruc.2008.05.003>
- Statistics: School and pupil numbers*. (2019). GOV.UK. <https://www.gov.uk/government/collections/statistics-school-and-pupil-numbers>
- Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods*, 8(4), 537–553. <https://doi.org/10.1002/jrsm.1260>

- Tseng, S.-S., & Yeh, H.-C. (2018). Integrating Reciprocal Teaching in an Online Environment with an Annotation Feature to Enhance Low-Achieving Students' English Reading Comprehension. *Interactive Learning Environments*, 26(6), 789–802. eric.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes / L. S. Vygotsky; edited by Michael Cole ... [et al.]; [translated from the Russian by Alexander R. Luria ... [et al.]]*. Harvard University Press.
- Wilson, D., B. (n.d.). *Practical Meta-Analysis Effect Size Calculator [Online calculator]*. Campbell Collaboration. Retrieved January 11, 2020, from <https://campbellcollaboration.org/research-resources/effect-size-calculator.html>
- Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Yeh, H.-C., Hung, H.-T., & Chiang, Y.-H. (2017). The Use of Online Annotations in Reading Instruction and Its Impact on Students' Reading Progress and Processes. *ReCALL*, 29(1), 22–38. eric.

Appendices

Appendix 1 – WoE A coding protocol example

Coding Protocol: Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children*, 71(2), 149–164. <https://doi.org/10.1177/001440290507100202>

Study: (Choo, Eng, & Ahmad, 2011)

Essential quality indicators

Quality Indicators for Describing Participants

Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?

- Yes N/A
- No No information provided

Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?

- Yes N/A
- No No information provided

Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?

- Yes N/A
- No No information provided

Quality Indicators for Implementation of the Intervention and Description of Comparison Conditions

Was the intervention clearly described and specified?

Yes N/A

No No information provided

Was the fidelity of implementation described and assessed?

Yes N/A

No No information provided

Was the nature of services provided in comparison conditions described?

Yes N/A

No No information provided

Quality Indicators for Outcome Measures

Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalized performance?

Yes N/A

No No information provided

Were outcomes for capturing the interventions effect measured at the appropriate times?

Yes N/A

No No information provided

Quality Indicators for Data Analysis

Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?

Yes N/A

No No information provided

Did the research report include not only inferential statistics but also effect size calculations?

Yes N/A

No No information provided

Desirable Quality Indicators

Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

Yes N/A

No No information provided

Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?

Yes N/A

No No information provided

Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

Yes N/A

No No information provided

Was evidence of the criterion-related validity and construct validity of the measures provided?

Yes N/A

No No information provided

Did the research team assess not only surface features of fidelity implementation (e.g., number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?

Yes N/A

No No information provided

Was any documentation of the nature of instruction or series provided in comparison conditions?

Yes N/A

No No information provided

Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

Yes N/A

No No information provided

Were results presented in a clear, coherent fashion?

Yes N/A

No

No information provided

Appendix 2 – Criteria for WoE B

Table 9

Weight of Evidence B (WoE B) Criteria

WOE B Rating	Criteria
3	Randomised controlled trial <ul style="list-style-type: none"> - Includes a control group - Measures taken pre and post intervention
2	Quasi-experimental designs or cohort studies <ul style="list-style-type: none"> - Includes a control group - Measures taken pre and post intervention
1	All other types of research design <ul style="list-style-type: none"> - No control group - Qualitative data taken - No pre and post intervention measures taken

Note – the rationale for WOE B is based on Petticrew and Roberts (2003) typologies of evidence where the types of research designs most appropriate for answering ‘effectiveness’ questions were considered.

Appendix 3 – WoE C Criteria

Weight of Evidence C (WoE C) Criteria

Criteria	Weighting and Descriptor	Rationale
Intervention method	3 – carried out face to face in a small group with an adult trained in RT methods using appropriate materials 2 – not carried out face to face but still uses RT principles OR carried out face to face and in a small group but RT principles are not the only method used 1 – unclear if RT principles are used or unclear exactly how the intervention is carried out	This keeps to RT principles and ensures that RT is used in a 'pure' form as designed by Palincsar and Brown (1984).
Intervention language	3 – person implementing the intervention uses English 2 – person implementing the intervention uses English and another language 1 – not carried out in English	The results will be more generalisable to UK educational settings where the intervention is likely to be carried out in English.
Participant level of English	3 – participants are clearly defined as learning English as a second or additional language and a quantifiable measure of their current English language ability is provided 2 – participants are clearly defined as learning English as a second or additional language, but no quantifiable measure of their current English language ability is provided 1 – level of English is not reported	The review question regards EAL learners – therefore their level of English is pertinent to the outcomes of the findings.
Setting of study	3 – study is carried out in a UK secondary school, sixth form or further education establishment 2 – study is carried out in a non-UK secondary school, sixth form or further education establishment equivalent 1 – study is carried out in a higher education establishment	The results will be more generalisable to UK educational settings.

Criteria	Weighting and Descriptor	Rationale
Country of study	3 – study is carried out in the UK 2 – study is carried out in a country where the official first language is English 1 – study is carried out in a country where the official first language is not English	The results will be more generalisable to UK educational settings.
Measures	3 – a standardised measure of reading comprehension is used 2 – a non-standardised measure of reading comprehension that is not researcher-developed is used. Includes researcher adapted measures. 1 – a researcher-developed measure of reading comprehension is used	It has been found that using researcher-developed measures of reading comprehension has led to higher effect sizes compared to standardised measures (Rosenshine & Meister, 1994). Therefore, to increase generalisability researcher-developed measures of reading comprehension are being weighted lower.

Appendix 4 – WoE C Results

Summary of Weight of Evidence C (WoE C) Ratings for Included Studies

Study	Intervention method	Intervention language	Participant level of English	Setting of study	Country of study	Measures	WoE C Rating
Choo et al. (2011)	3	3	2	2	1	2	2.2
Dabarera et al. (2014)	3	3	3	2	1	2	2.3
Fung et al. (2002)	3	2	3	2	2	3	2.5
Huang and Yang (2015)	2	3	3	1	1	2	2
Klinger and Vaughn (1996)	3	3	3	2	2	3	2.7
Soonthornmanee (2002)	1	3	3	1	1	1	1.7
Tseng and Yeh (2018)	2	3	3	1	1	3	2.2
Yeh et al. (2017)	2	3	3	1	1	3	2.2

Note – final WoE C Rating was calculated by averaging across all WoE C criteria scores and rounding to one decimal place.

Appendix 5 – Details of excluded articles and reason for exclusion

Articles excluded at title and abstract screening stage

Study	Exclusion criteria met
Adams, P. E. (1995). Teaching ‘Romeo and Juliet’ in the Nontracked English Classroom. <i>Journal of Reading</i> , 38(6), 424–432. Retrieved from eric.	2
Ahmadi, M. R., & Gilakjani, A. P. (2012). Reciprocal Teaching Strategies and Their Impacts on English Reading Comprehension. <i>Theory and Practice in Language Studies</i> , 2(10), 2053–2060.	2
Armbrister, A. L. (2010). Non-Native Speakers Reach Higher Ground: A Study of Reciprocal Teaching’s Effects on English Language Learners.	2
Brown, J. D. (Series Ed.). (2013). <i>New Ways of Classroom Assessment</i> . Revised.	2
Buckley, P., Lasser, C. J., Scornavacco, K., Klingner, J., & Boardman, A. (2013). Strengthening comprehension with Collaborative Strategic Reading. <i>Better: Evidence-Based Education</i> , 5(3), 12–13.	2
Casanave, C. P. (1988). Comprehension Monitoring in ESL Reading: A Neglected Essential. <i>TESOL Quarterly</i> , 22(2), 283–302.	2
Casey, J. Elizabeth, Mireles, S. V., Vilorio, M. de L., & Garza, E. (2018). Literacy & Arts Integration in Science: Engaging English Language Learners in a Lesson on Mixtures and Solutions. <i>Texas Journal of Literacy Education</i> , 6(1), 51–69.	2
Casey, Jane Elizabeth. (2012). A formative experiment to increase English language learners’ awareness and use of metacognitive strategies through reciprocal teaching: Pushing toward an end to silence in the classroom. <i>Dissertation Abstracts International Section A: Humanities and Social Sciences</i> , 72(8-A), 2688.	1
Cotterall, S. (1990). Developing Reading Strategies through Small-Group Interaction. (No. 0033–6882; pp. 55–69).	2
Erten, I. H., & Karakas, M. (2007). Understanding the Divergent Influences of Reading Activities on the Comprehension of Short Stories. <i>Reading Matrix: An International Online Journal</i> , 7(3), 113–133.	3
Frances, S. M., & Eckart, J. A. (1992). The Effects of Reciprocal Teaching on Comprehension.	4
Garderen, D. v. (2004). Focus on inclusion reciprocal teaching as a comprehension strategy for understanding mathematical word problems. <i>Reading and Writing Quarterly</i> , 20(2), 225-229. doi:10.1080/10573560490272702	3
Ghorbani, M. R., Gangeraj, A. A., & Alavi, S. Z. (2013). Reciprocal Teaching of Comprehension Strategies Improves EFL Learners’ Writing Ability. <i>Current Issues in Education</i> , 16(1), 1–13.	5
Gilbert, F. (2018). Riding the Reciprocal Teaching Bus. A Teacher’s Reflections on Nurturing Collaborative Learning in a School Culture Obsessed by Results. <i>Changing English: Studies in Culture and Education</i> , 25(2), 146–162.	2

Study	Exclusion criteria met
Gomez, L., & And Others. (1995). Natural Assessment of Oral Language Growth of Limited English Proficient Students in Paired Reciprocal Learning (pp. 1–21).	5
Guccione, L. M. (2011). Integrating Literacy and Inquiry for English Learners. <i>Reading Teacher</i> , 64(8), 567–577.	2
Hancock, J. W. (2012). The Effect of Reciprocal Teaching Plus across Three Content Areas on Reading Comprehension in Seventh and Eighth Grade Students (ProQuest LLC).	4
Harper, C., Cook, K., & James, C. K. (2010). Content-language integrated approaches for teachers of EAL learners: Examples of reciprocal teaching. <i>English as an additional language: Approaches to teaching linguistic minority students</i> (pp. 75-96) doi:10.4135/9781446251454.n6	2
Hosenfeld, C., & And Others. (1993). Activities and Materials for Implementing Adapted Versions of Reciprocal Teaching in Beginning, Intermediate, and Advanced Levels of Instruction in English, Spanish, and French as a Second/Foreign Language.	2
Hoyt, L. (2002). <i>Make It Real: Strategies for Success with Informational Texts</i> . 1–320.	2
Hsu, C., Yang, F. - O., & Wu, W. - V. (2013). An e-learning tool for blended reciprocal teaching on english textbook for EFL technology-majored students. Paper presented at the Proceedings of the 21st International Conference on Computers in Education, ICCE 2013, 824-826.	1
Kang, H. (1998). The Importance of 'Metacognition' in ESL Reading Instruction. <i>Journal of Japan-Korea Association of Applied Linguistics</i> , 1, 25–32.	2
Lawrence, L. J. (2007). Cognitive and Metacognitive Reading Strategies Revisited: Implications for Instruction. <i>Reading Matrix: An International Online Journal</i> , 7(3), 55–71.	2
Liu, A., & Bu, Y. (2016). Reciprocal Learning Strategy in CALL Environment: A Case Study of EFL Teaching at X University in Shanghai. <i>Universal Journal of Educational Research</i> , 4(5), 1059–1070.	2
Miami-Dade County Public Schools, R. S. (2008). Miami-Dade County Public Schools Statistical Abstract 2007-2008. Research Services, Miami-Dade County Public Schools.	2
Miller, L. D., & Perkins, K. (1989). <i>ESL Reading Comprehension Instruction</i> . 1–27.	2
Mullen-O'Leary, C. (2016). English language learners' reading comprehension: Advance organizers and comprehension monitoring. <i>Dissertation Abstracts International Section A: Humanities and Social Sciences</i> , 76(10-A(E)),	1
Oczkus, L. D. (2018). <i>Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension</i> , 3rd Edition.	2
Ramos, J. A. (2012). Ameliorating the English Reading Comprehension of Spanish-Speaking ELLs through a Reciprocal Teaching Intervention.	1
Soto, L. D. (1989). Enhancing the Written Medium for Culturally Diverse Learners via Reciprocal Interaction. <i>Urban Review</i> , 21(3), 145–149.	2

Study	Exclusion
	criteria met
Stäbler, F., Dumont, H., Becker, M., & Baumert, J. (2017). What Happens to the Fish's Achievement in a Little Pond? A Simultaneous Analysis of Class-Average Achievement Effects on Achievement and Academic Self-Concept. <i>Journal of Educational Psychology</i> , 109(2), 191–207.	3
Suhor, C. (1988). Beyond 'Trends' in English and Language Arts Instruction.	3
Sun, L.-E. (2012). A study of the effects of reciprocal teaching as a reading strategy instruction on metacognitive awareness, self-efficacy, and English reading comprehension of EFL junior high school students. <i>Dissertation Abstracts International Section A: Humanities and Social Sciences</i> , 72(12-A), 4427.	1
Van Garderen, D. (2004). Reciprocal Teaching as a Comprehension Strategy for Understanding Mathematical Word Problems. <i>Reading and Writing Quarterly</i> , 20(2), 225–229.	2
Wayman, M. M., McMaster, K. L., Saenz, L. M., & Watson, J. A. (2010). Using Curriculum-Based Measurement to Monitor Secondary English Language Learners' Responsiveness to Peer-Mediated Reading Instruction. <i>Reading & Writing Quarterly</i> , 26(4), 308–332.	2
Williams, J. A. (2010). Taking on the Role of Questioner: Revisiting Reciprocal Teaching. <i>Reading Teacher</i> , 64(4), 278–281.	2