<u>How effective is metacognitive instruction at improving the word problem-solving of children who are low-achievers in maths?</u>

Summary

Metacognitive knowledge and skills are powerful predictors of academic outcomes (Wang et al., 1990) but are often lacking in children who are low-achievers in maths (Miller & Mercer, 1997). Metacognitive instruction seeks to address this deficit (Veenman, 2015). Interventions address strategy knowledge (plan-monitor-evaluate), task knowledge (when and why to apply strategies), and person knowledge (strengths, weaknesses, and motivation), and provide opportunity to practise (Flavell, 1979; Livingston, 1996; Pintrich, 2002). This review sought to evaluate the effect of metacognition interventions on mathematical word problem-solving. A systematic literature search was conducted, identifying seven studies for review. A meta-analysis showed a large combined effect size (g = 1.39) when comparing intervention to comparison participants. This, in combination with sufficient methodological quality among the reviewed studies, suggests metacognitive instruction can be recommended as evidence-based practice (Gersten et al., 2005). Recommendations for educational psychology practice, limitations of the review, and recommendations for future research are discussed.

Introduction

Metacognition

Metacognition refers to knowledge about, and regulation of, cognition (Schraw, 1998). While cognitive skills are used to perform tasks (such as multiplication), metacognitive skills are used to decide how to perform tasks and to evaluate performance (Garner, 1987). Psychological research on metacognition began in earnest in the 1970s (Gleitman et al., 1972). It has since been acknowledged as a concept of profound psychological importance, being incorporated into a revision of Bloom's Taxonomy of Learning as a fourth dimension of knowledge (Krathwohl, 2002).

There are two prominent theoretical models of metacognition. Flavell (1979) distinguishes four components: knowledge, experience, goals, and actions (Figure 1). *Knowledge* comprises three sub-components: person, task, and strategy. 'Person knowledge' involves awareness of oneself and others as cognitive processors, including strengths, weaknesses, and motivation. 'Task knowledge' involves awareness of how to manage cognitive enterprises, including implications of task difficulty and situational norms for strategy selection (Pintrich, 2002). 'Strategy knowledge' involves awareness of ways of effectively achieving cognitive goals, including planning, monitoring, evaluating, information-acquisition strategies (e.g. mnemonics), and problem-solving heuristics. Metacognitive *experiences* are conscious feelings accompanying cognitive enterprises, such as being aware that one does not understand something. *Goals* refer to awareness of task objectives and *actions* refer to strategies or behaviours employed to achieve goals.

An alternative model (Schraw, 1998) distinguishes two components: knowledge and regulation (Figure 2). *Knowledge* comprises three sub-components: declarative, procedural, and conditional. 'Declarative knowledge' involves awareness about oneself and factors influencing performance. 'Procedural knowledge' involves awareness of effective strategies and heuristics to complete tasks. 'Conditional knowledge' involves awareness of when and why to use declarative and procedural knowledge, such as allocating resources and selecting strategies. *Regulation* is the active employment of knowledge before, during, and after a task to plan, monitor, and evaluate learning and performance.

Conceptual similarities across the models include self-awareness of strengths, weaknesses, and motivation; knowledge of how and when to use cognitive strategies; and the planning, monitoring and evaluating sequence. Psychometric evidence supports the parsimony of a two-factor model (Schraw & Dennison, 1994). Schraw and Dennision conducted unrestricted factor analysis of a 52-item metacognitive inventory. This produced an unreliable six-factor solution but it did not map onto the six conceptual sub-components. Restricted factor analysis, however, strongly supported a two-factor solution (knowledge and regulation), with high internal consistency ($\alpha = .91$) on each factor and 44 items loading unambiguously onto a single factor. Furthermore, the factors contributed separately to performance on a reading comprehension test, suggesting the need to develop both metacognitive knowledge and regulation skills for optimal outcomes.

Figure 1. A Four-Component Model of Metacognition, Adapted From Flavell (1979).

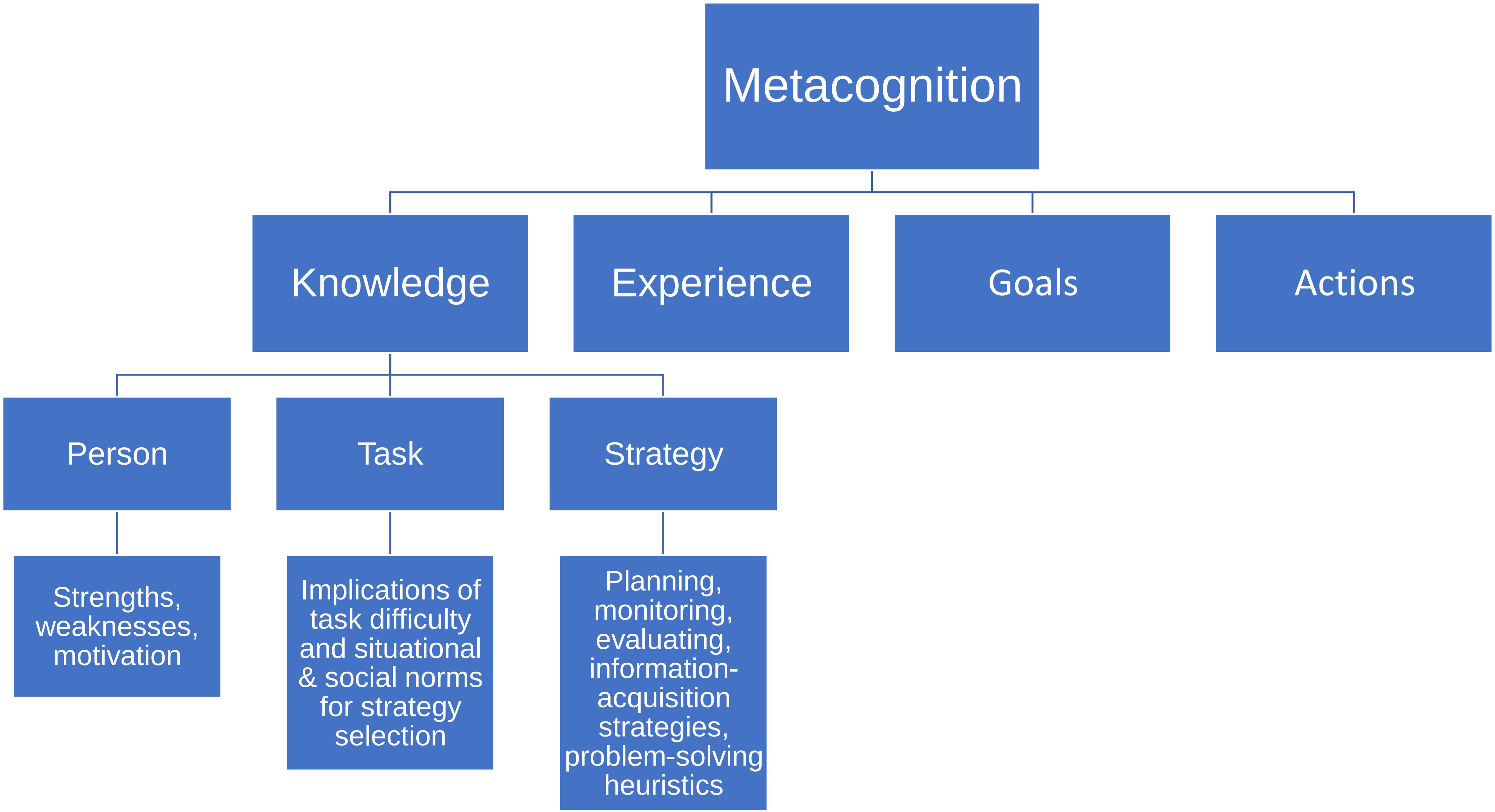
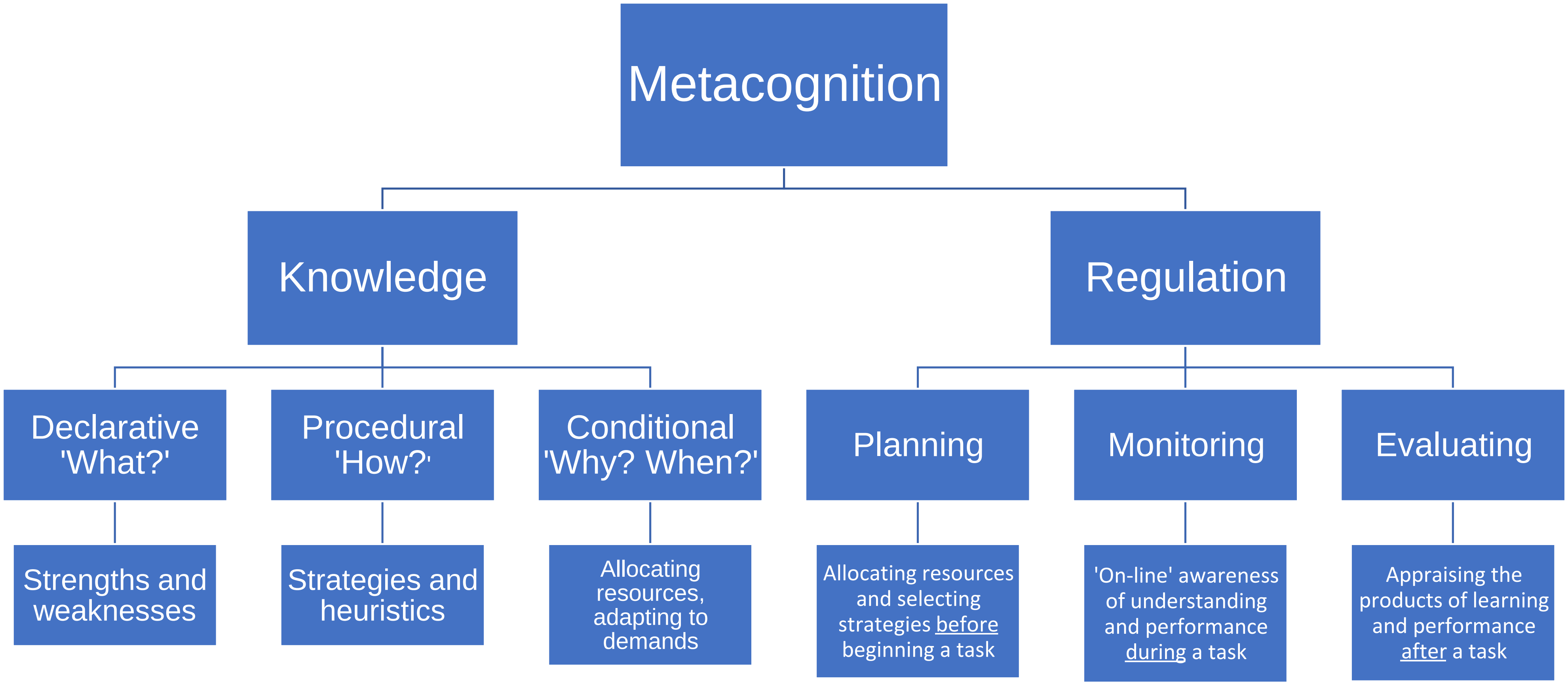


Figure 2. A Two-Component Model of Metacognition, Adapted From Schraw (1998).



Teaching metacognition

According to the Education Endowment Foundation, teaching metacognition in schools has high impact for very low cost, based on extensive evidence (Quigley et al., 2018). Metacognition was identified as the single most important predictor of learning outcomes, above classroom management, student/teacher interactions and 27 further variables (Wang et al., 1990).

Metacognitive skills may be acquired as part of typical development, emerging around age six and increasing in frequency and quality from age eight (Veenman & Spaans, 2005). Metacognition develops first as a domain-specific attribute but, with growing proficiency around age 14, generalises and promotes learning transfer between contexts (Schraw, 1998; Veenman & Spaans, 2005). Metacognitive skills developed in one subject-area benefit individuals in other subject-areas and in life beyond formal education (Pintrich, 2002).

However, some children require explicit instruction to acquire metacognitive skills (Veenman, 2015). An important area of learning that requires metacognition and has everyday relevance is mathematical problem-solving (Montague, 1997). This is a foundational skill for school attainment and is vital for everyday tasks such as grocery shopping. Most students struggle with metacognitive components such as assessing their ability, selecting appropriate strategies, organising information, monitoring, and evaluating outcomes (De Corte et al., 2000; Miller & Mercer, 1997).

A prominent intervention addressing metacognitive skills in mathematical problem-solving is Cognitive Strategy Instruction (CSI) (Montague et al., 2011). CSI combines metacognitive knowledge, regulation, and cognitive strategies in a sequential problem-solving model. Children memorise seven cognitive strategies (read, paraphrase, visualise, hypothesise, estimate, compute, check) and perform metacognitive strategies (say, ask, check) at each step to ensure they have completed the step comprehensively.

A systematic literature review of CSI identified five single-subject and two group-experimental designs (Montague & Dietz, 2009). Despite consistent evidence of the effectiveness of CSI at improving problem-solving, the review concluded that findings did not meet the methodological criteria for evidence-based practice. It suggested future research needed more stringent experimental designs. In light of this, the current review broadened the scope of intervention, seeking any study employing metacognitive instruction, but narrowed the scope of experimental design, seeking only studies with pre-/post-data on intervention and comparison groups. This review aims to answer the question:

How effective is metacognitive instruction at improving the word problem-solving of children who are low-achievers in maths?

Critical review of the evidence base

Literature search

A systematic literature search was conducted on 15 December 2019 using three online databases: Web of Science, Educational Resources Information Center (ERIC), and PsycINFO. Search terms related to variations on ‘metacognition’ and variations on ‘maths learning difficulties’. Database searches yielded 248 results. Following removal of 31 duplicates, 217 articles underwent title and abstract screening to determine eligibility for inclusion in the review. Inclusion criteria specified studies must have at least one intervention and comparison group, an intervention involving metacognitive instruction, a quantitative outcome measure of mathematical problem-solving, and take place in schools with children aged 5-16 who were low-achievers in maths. One hundred and seventy-seven articles were excluded, leaving 40 articles for full text screening. Ten additional articles were identified through ancestral and citation searching and screened at full text. Forty-three studies were excluded, leaving seven studies eligible for review.

The seven included studies were critically appraised using the Weight of Evidence (WoE) framework (Gough, 2007). Dimensions considered were methodological quality (WoE A), methodological relevance (WoE B), and topic relevance (WoE C). WoE A was a generic judgment of the quality of the research design including participant description, intervention implementation, outcome measures, and data analysis. A published coding protocol was used to assess WoE A (Gersten et al., 2005). WoE B and C were judgments relating to the review question, using coding protocols developed by the author. WoE D is the average of WoE A, B, and C. A summary of key information about the included studies is provided in Table 1.

Table 1. Key Information about the Seven Included Studies.

Study	Research design	Main findings	Effect Size Description	Effect Size	WoE
Chung & Tam (2005) Country: China (Hong Kong) Sample size: 30 (8 girls) Age: *M* = 10:4 years Setting: Special school	Intervention group [5-step model based on CSI] Comparison group 1 [taught to visualise problems] Comparison group 2 [extra regular teaching] All *n* = 10 Intervention delivered in groups of 10 with the first author during regularly scheduled resource classes in five 50 minute sessions	Intervention (I) and Comparison 1 (C1) scored significantly higher at post-test and follow-up than Comparison 2 (C2) but did not differ from each other Post-test scores were maintained at follow-up by both I and C1 but not by C2	Calculation: Post-test, comparison vs intervention, *M* & *SD* I vs C1 I vs C2 C1 vs C2	-0.09 (small) 1.17 (large) 1.40 (large)	WoE A: 1 Low WoE B: 2 Medium WoE C: 1.8 Medium WoE D: 1.6 Medium

Study	Research design	Main findings	Effect Size Description	Effect Size	WoE
Fuchs et al. (2003) <u>Country:</u> USA <u>Sample size:</u> 395 <u>Age:</u> 3rd grade <u>Setting:</u> Six mainstream schools	Intervention group (*n* = 137) [word problem-solving practice plus Self-regulated learning] Comparison group 1 (*n* = 138) [word problem-solving practice only] Comparison group 2 (*n* = 120) [no extra teaching] Intervention delivered by research assistants and teachers in 32 sessions of 30-40 minutes	Participants completed 10 *immediate transfer* word problems (four problem structures with novel cover stories), seven *near transfer* problems (four problem structures, novel cover stories, one superficial problem feature varied), one *far transfer* problem (all structures embedded in a real-life context, with all superficial features varied and elements of novelty) Immediate transfer - C2 improvement was less than C1, which in turn was less than I; this was found across low, average and high achieving participants Near transfer – average and low achieving participants in C1 and I improved more than C2 but did not differ from each other Far transfer – C2 improvement was less than C1, which in turn was less than I; this was found across low, average and high achieving participants	<u>Calculation:</u> Pre- vs post-test, comparison vs intervention, *M* & *SD* (low-achievers only) *Immediate transfer* I vs C1 I vs C2 C1 vs C2 *Near transfer* I vs C1 I vs C2 C1 vs C2 *Far transfer* I vs C1 I vs C2 C1 vs C2	 0.33 (small) 2.68 (large) 1.83 (large) 0.35 (small) 2.18 (large) 1.24 (large) 0.21 (small) 1.17 (large) 0.69 (medium)	WoE A: 3 High WoE B: 2.5 High WoE C: 2 Medium WoE D: 2.5 High

Study	Research design	Main findings	Effect Size Description	Effect Size	WoE
Kajamies et al. (2010) Country: Finland Sample size: 429; 24 main participants (12 girls) and a large extra comparison group Age: 4th grade (*M* = 10:4 years) Setting: Twelve mainstream schools	Intervention group (*n* = 8) [computer game teaching a 6-step problem-solving model] Comparison group 1 (*n* = 8) [no extra teaching] Comparison group 2 (*n* = 8) [reading comprehension intervention] Comparison group 3 (*n* = 405) [no extra teaching but represented the range of attainment levels in the schools] Intervention delivered in groups of two with the first author in a quiet room at school in 14 sessions of 45 minutes	All four groups increased their scores at both post-test and follow-up I participants' scores increased from pre- to post-test significantly more than C3 participants' scores At follow-up, I participants' scores no longer significantly differed from C3 participants' scores (but C1 and C2 were still lower than both) C1 and C2 participants' scores did not increase at a differential rate	Calculation: Pre- vs post-test, comparison vs intervention, *M* & *SD* I vs C1 I vs C2	0.74 (medium -large) 0.67 (medium)	WoE A: 1 Low WoE B: 1.75 Medium WoE C: 2.4 Medium WoE D: 1.72 Medium

Study	Research design	Main findings	Effect Size Description	Effect Size	WoE
Pennequin et al. (2010) Country: France Sample size: 48 (25 girls) Age: 3rd grade (*M* = 8:10 years) Setting: Mainstream school	Intervention group [based on the Strategy Evaluation Matrix (Schraw, 1998)] Comparison group [extra regular teaching] Both *n* = 24 Intervention delivered in groups of 6 by a research assistant in 5 sessions of 60 minutes	Both normal- and low-achievers in the I group had higher post-test scores but neither C group had higher post-test scores I low-achievers improved more than normal-achievers At post-test there was no longer a difference in scores between I low- and normal-achievers On a measure of metacognitive knowledge, only low-achievers had increased post-test scores	Calculation: Pre- vs post-test, comparison vs intervention, *F* value of Pre/post x Group interaction I vs C	1.21 (large)	WoE A: 2 Medium WoE B: 2.5 High WoE C: 1.2 Low WoE D: 1.9 Medium
Teong (2003) Country: Singapore Sample size: 40 Age: 11-12 years Setting: Mainstream school	Intervention group [computer game teaching a 5-step problem-solving model] Comparison group [word problem-solving practice only] Intervention delivered in 4 sessions of 60 minutes	I scores increased from pre- to post-test and increased further at follow-up Post-test I scores were significantly higher than C scores	Calculation: Pre- vs post-test, comparison vs intervention, *M* & *SD* I vs C	0.91 (large)	WoE A: 0 Low WoE B: 2.25 Medium WoE C: 1.2 Low WoE D: 1.15 Low

Study	Research design	Main findings	Effect Size Description	Effect Size	WoE
Wang et al. (2019) <u>Country:</u> USA <u>Sample size:</u> 69 (36 girls) <u>Age:</u> 3rd grade <u>Setting:</u> Six Mainstream schools	Intervention group (*n* = 23) [fractions and self-regulation teaching] Comparison group 1 (*n* = 24) [fractions teaching only] Comparison group 2 (*n* = 26) [no extra teaching] Intervention delivered by trained tutors in 39 sessions of 35 minutes	I post-test scores were higher than C2 C1 post-test scores were also higher than C2 but there was a moderation effect of pre-test scores; C1 participants with higher pre-test scores responded more adequately to the fractions teaching than those with lower pre-test scores This moderation effect was not apparent for I	<u>Calculation:</u> Post-test, comparison vs intervention, *M* & *SD* I vs C1 I vs C2 C1 vs C2	-0.04 (small) 1.00 (large) 0.91 (large)	WoE A: 3 High WoE B: 2.5 High WoE C: 1.8 Medium WoE D: 2.43 Medium-High

Study	Research design	Main findings	Effect Size Description	Effect Size	WoE
Zhu (2015) Country: China Sample size: 150 (63 girls) Age: 4th grade Setting: Mainstream school	Intervention group [7-step problem-solving model based on CSI] Comparison group [extra regular teaching] Both *n* = 75 Within Intervention and Comparison, participants were divided into 4 ability groups: 1 (maths difficulties only), 2 (maths and reading difficulties), 3 (average achieving), 4 (high achieving) Intervention delivered by teachers in 16 sessions of 40 minutes	All four intervention groups outperformed their respective comparison groups I groups all had significant improvement from pre- to post-test C low-achievers showed no or little response to regular teaching I participants with only low maths scores benefitted more from the intervention than those with low maths and reading scores Group 4 (high-achievers) benefitted less from the intervention than the other three groups	Calculation: Pre- vs post-test, comparison vs intervention, *M* & *SD* I vs C (Maths Difficulties (MD) only) I vs C (MD & Reading Difficulties) I vs C (Average Achieving) I vs C (High Achieving)	1.82 (large) 0.99 (large) 1.48 (large) 0.72 (medium -large)	WoE A: 3 High WoE B: 2.5 High WoE C: 1.6 Medium WoE D: 2.37 Medium-High

Participants

In total, 1161 participants took part in the reviewed studies, ranging from age 8 to 12 years. There was substantial variation in sample size, from 30 to 429. From available data, sex representation was roughly equal with 45% female participants (144/321). Studies took place in China, Finland, France, Singapore, and the USA. There was thus substantial heterogeneity in cultural background and educational systems among participants. This has positive implications for the generalisation of findings and potentially allows for cross-cultural analysis. A potential drawback for EPs working in the UK is the absence of UK-based evidence.

Research design

All studies used an experimental design with pre-/post-testing and intervention/comparison groups. There was a variety of comparison groups. Fuchs et al. (2003), Teong (2003), and Wang et al. (2019) isolated the effect of metacognitive instruction by including a comparison group which received equivalent teaching with equivalent delivery parameters minus the metacognitive components. Chung and Tam (2005), Pennequin et al. (2010), and Zhu (2015) had groups receiving additional regular maths teaching, while alternative interventions were provided by Chung and Tam (2005) (taught to visualise problems) and Kajamies et al. (2010) (reading comprehension instruction). These controlled for attention effects (McCarney et al., 2007) but were less able to isolate the effect of metacognitive instruction. Fuchs et al. (2003), Kajamies et al. (2010), and Wang et al. (2019) had groups receiving no additional teaching, representing the starkest contrast with participants receiving metacognitive instruction. Three studies included a second comparison group receiving regular teaching (Chung & Tam, 2005; Fuchs et al., 2003; Wang et al., 2019), facilitating comparison of the metacognition intervention with both an alternative intervention and regular teaching.

Studies which took follow-up measures (Chung & Tam, 2005; Kajamies et al., 2010; Teong, 2003) were rated higher in WoE C because this illustrated whether intervention benefits were maintained. Studies which measured other attributes of metacognition in addition to problem-solving scores (Fuchs et al., 2003; Kajamies et al., 2010; Pennequin et al., 2010; Teong, 2003) were rated higher in WoE C because this provided a more holistic picture and indicated whether participants could generalise learning. Studies which took ‘far-transfer’ measures (word-problems structured differently to those practised during intervention) were also rated higher in WoE C (Chung & Tam, 2005; Fuchs et al., 2003) because this indicated whether participants could apply learning in a novel mathematical context.

Intervention

Interventions ranged from 4 to 22.75 hours of total delivery time (M = 10.72, SD = 7.29) in 4 to 39 sessions lasting between 30 and 60 minutes between 1 and 3 times per week, indicating substantial heterogeneity. Interventions were delivered by researchers or research-assistants apart from those implemented by Fuchs et al. (2003) and Zhu (2015), who trained teachers. This contributed to external validity, showing teachers with two days’ training could deliver interventions. Teong (2003) did not state who delivered the intervention, hindering replicability.

Content and procedures of metacognitive instruction differed. No study taught all areas of metacognitive knowledge (Flavell, 1979). All studies combined teacher instruction with independent practice. Four studies provided participants with sequential problem-solving models; three (Chung & Tam, 2005; Kajamies et al., 2010; Zhu, 2015) were derived from CSI and one (Teong, 2003) was researcher-developed but similar. These studies focused on strategy knowledge, particularly the plan-monitor-evaluate sequence. Only Kajamies et al. (2010) addressed person knowledge – engaging participants in peer discussion – and task knowledge – deciding which strategy was appropriate for each task (also addressed by Teong, 2003). Pennequin et al. (2010) adopted a similar strategic focus without provision of a problem-solving model. Person and task knowledge were addressed as by Kajamies et al. (2010) but participants were not given mathematical problem-solving teaching.

Fuchs et al. (2003) and Wang et al. (2019) adopted person-focused teaching, labelling their interventions ‘self-regulated learning’. There was a focus on analysing participants’ strengths and weaknesses through goal-setting, marking and evaluating work, and tracking progress. Fuchs et al. (2003) were the only researchers to discuss with participants how they had transferred learning to other subjects or areas outside of school. Since metacognitive skills are potentially domain-general (Schraw, 1998), applying skills beyond the intervention context is likely a helpful learning process.

Findings

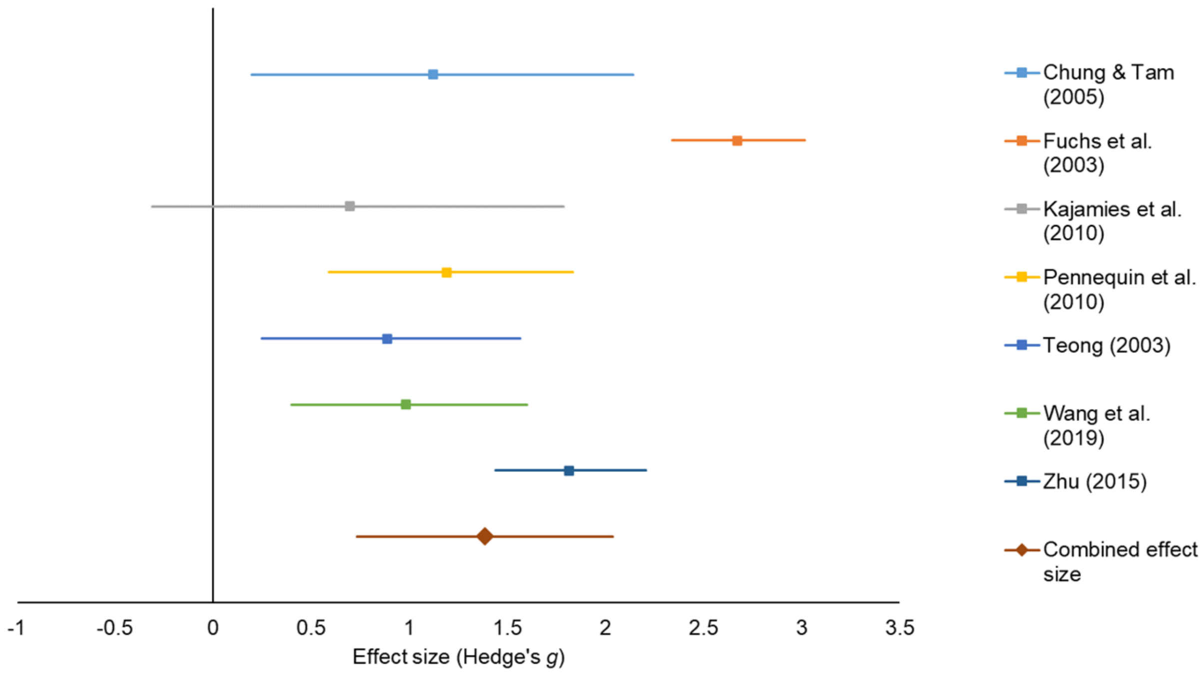
The effect size calculated for all studies was the standardised mean difference (Hedge's *g*). Where possible, this was calculated by the author as the difference between intervention and comparison improvement (post-test minus pre-test) means divided by the pooled standard deviation of pre-test means (Morris, 2008). If there were insufficient data, only post-test means were used. Pennequin et al. (2010) provided no descriptive statistics so effect sizes were calculated using the *F*-statistic of the interaction between pre-/post-scores and intervention/comparison with the Campbell Collaboration online calculator (Wilson, n.d.).

A meta-analysis was conducted to assess the overall effect of metacognitive instruction on problem-solving (Figure 3). The comparison groups in the meta-analysis were those which provided the biggest experimental contrast from each study (a group receiving no additional teaching or extra regular teaching). While this does not allow consideration of the most effective way to deliver metacognitive instruction, it gives an average baseline figure against which future meta-analyses could compare (Law et al., 2004). The meta-analysis was conducted with a random-effects model using *Meta-Essentials* software (Suurmond et al., 2017).

The combined effect size across seven studies was $g = 1.39$, 95% CI [0.73, 2.04]). This can be described as large, with a medium-large lower confidence interval. This statistical evidence is supported by WoE D ratings. Fuchs et al. (2003) were rated High and had the largest effect size, while Wang et al. (2019) and Zhu (2015) were rated Medium-High and had large effect sizes. The average WoE D rating across studies was 1.95 ($SD = .5$). It was hypothesised that effect size may be related to length of intervention, but this correlation was unclear and non-significant, $r = .36$, 95% CI [-.63, .997], $p = .426$.

A second meta-analysis was conducted using comparison groups that received the same instruction as intervention groups minus the metacognitive components. This attempted to isolate the effect of metacognitive instruction from all other effects of intervention including increased attention from teachers, problem-solving practice, and the novel experience of taking part in research. A more conservative picture emerged with a combined effect size of $g = 0.36$, 95% CI [-0.65, 1.37], which is small. Since the lower CI crosses zero, there is poor statistical evidence that metacognitive instruction provided benefit beyond the other components of intervention.

Figure 3. Forest Plot of Effect Sizes (95% Confidence Intervals) from the First Meta-Analysis.



Discussion

This review evaluated whether metacognitive instruction improved word problem-solving of children who were low-achievers in maths. Seven studies met the inclusion criteria, with one receiving a High WoE D rating, two Medium-High, three Medium, and one Low.

Given the combined evidence of statistical effect, methodological quality, and methodological and topical relevance, it can be concluded that interventions incorporating metacognitive instruction had a considerable effect on problem-solving. Gersten et al. (2005) suggested, for an intervention to be evidence-based practice, there should be two studies with High WoE A (three were found in this review) and a combined effect size significantly greater than zero. This review supports the claim for maths interventions incorporating metacognitive instruction as evidence-based practice.

Evidence for the unique contribution of metacognitive instruction above other intervention components is equivocal. Based on the second meta-analysis, it cannot be confidently concluded that there was an effect on problem-solving. However, studies with follow-up measures (Chung & Tam, 2005; Kajamies et al., 2010; Teong, 2003) found intervention participants maintained gains to a greater degree than comparison participants. Metacognitive instruction may promote longer-term learning but it is difficult to assess given the lack of studies with follow-up measures and strong methodologies. Information from secondary outcomes is potentially enlightening. Fuchs et al. (2003) found through a questionnaire that intervention participants self-rated as having higher self-efficacy (d = 0.92) and higher goal orientation and self-monitoring (d = 1.2) than comparison participants who had the same intervention minus metacognitive components. Wang et al. (2019) found through a distal measure of general fraction tasks that intervention participants scored higher (d = 0.44) than comparison participants who had the same intervention minus metacognitive components. These findings tentatively support the theoretical claim that metacognitive knowledge and skills generalise beyond domains (Schraw, 1998), which may be a unique, additional contribution to conventional interventions.

Recommendations for practice

When considering the appropriateness of an intervention for educational psychology practice, evaluation of generalisability is key. The two studies which trained teachers to implement interventions (Fuchs et al., 2003; Zhu, 2015) had High or Medium-High WoE D ratings and large effect sizes, suggesting teacher delivery is feasible. Furthermore, three studies sampled from multiple mainstream or special schools, suggesting results generalised across settings. In the absence of a commercial intervention package, the only cost of metacognitive instruction is teacher training, either in CSI or general metacognitive principles. This is likely to have significant returns because teachers could utilise knowledge in classrooms and interventions.

In terms of participant characteristics, generalisability is less clear. Results did replicate across culturally disparate populations with different school structures. However, no studies took place in the UK. It may be inferred from evidence of cross-cultural replication that similar results would be found with a UK sample but this cannot be assumed.

Overall, given the substantial benefits of metacognitive instruction for children who are low-achievers in maths, and the simplicity and low cost of its implementation, it should be recommended by educational psychologists.

Limitations of the review

It could be argued this review's inclusion criteria permitted studies which taught cognitive as well as metacognitive strategies. However, given the domain-specific origins of metacognition (Schraw, 1998) it would seem conceptually and developmentally inconsistent to teach metacognition in isolation without any relevance to a particular subject, particularly for young learners who are struggling. Therefore, a review of studies which only taught metacognition would have had weaker external validity for educational psychology practice even if it provided stronger theoretical evidence. Furthermore, an attempt was made to isolate the effect of metacognition through a secondary meta-analysis.

Recommendations for future research

Previous authors have noted the need to identify which components of metacognition are important in facilitating change (Dowker, 2017). While this review did not address this issue, it did illustrate a dual focus in the literature on either strategy knowledge or self-regulation. Future research could explore the differential effect of these focal points and whether effects are additive. Given the apparent lack of correlation between total intervention length and effect size, future research could examine which parameters contribute to effective interventions such as session frequency, session length, and expertise of people delivering interventions. Finally, future studies should include follow-up measures and secondary outcomes to evaluate maintenance and generalisability. This is important given the domain-general nature of metacognition (Schraw, 1998) and its potential cross-curricular impact.

References

Chung, K. K. H., & Tam, Y. H. (2005). Effects of cognitive-based instruction on mathematical problem solving by learners with mild intellectual disabilities. *Journal of Intellectual and Developmental Disability*, *30*(4), 207–216. https://doi.org/10.1080/13668250500349409

De Corte, E., Verschaffel, L., & Eynde, P. O. (2000). Self-regulation: A characteristic and a goal of mathematics education. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 687–726). Academic Press. https://doi.org/http://dx.doi.org/10.1016/B978-012109890-2/50050-0

Dowker, A. (2017). Interventions for primary school children with difficulties in mathematics. In J. Sarama, D. H. Clements, C. Germeroth, & C. DayHess (Eds.), *Development of early childhood mathematics education* (Vol. 53, pp. 255–287). https://doi.org/10.1016/bs.acdb.2017.04.004

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1002/bit.23191

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003). Enhancing third-grade students' mathematical problem solving with self-regulated learning strategies. *Journal of Educational Psychology*, *95*(2), 306–315. https://doi.org/10.1037/0022-0663.95.2.306

Garner, R. (1987). *Metacognition and reading comprehension*. Ablex Publishing.

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

Gleitman, L. R., Gleitman, H., & Shipley, E. F. (1972). The emergence of the child as grammarian. *Cognition*, *1*(2–3), 137–164. https://doi.org/10.1016/0010-0277(72)90016-9

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

Kajamies, A., Vauras, M., & Kinnunen, R. (2010). Instructing low-achievers in mathematical word problem solving. *Scandinavian Journal Of Educational Research*, *54*(4), 335–355. https://doi.org/10.1080/00313831.2010.493341

Krathwohl, D. R. (2002). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. *Theory into Practice*, *41*(4), 212–218.

Law, J., Garrett, Z., & Nye, C. (2004). The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *47*(4), 924–943. https://doi.org/10.1044/1092-4388(2004/069)

Livingston, J. A. (1996). *Effects of metacognitive instruction on strategy use of college students [Unpublished manuscript]*. State University of New York at Buffalo.

McCarney, R., Warner, J., Iliffe, S., Van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: A randomised, controlled trial. *BMC Medical Research Methodology*, *7*. https://doi.org/10.1186/1471-2288-7-30

Miller, S. P., & Mercer, C. D. (1997). Educational aspects of mathematics disabilities. *Journal of Learning Disabilities*, *30*(1), 47–56.

Montague, M. (1997). Cognitive strategy instruction in mathematics for students with learning disabilities. *Journal of Learning Disabilities*, *30*(2), 164–177. https://doi.org/10.1177/002221949703000204

Montague, M., & Dietz, S. (2009). Evaluating the evidence base for cognitive strategy instruction and mathematical problem solving. *Exceptional Children*, *75*(3), 285–302. https://doi.org/10.1177/001440290907500302

Montague, M., Enders, C., & Dietz, S. (2011). Effects of Cognitive Strategy Instruction on Math Problem Solving of Middle School Students With Learning Disabilities. *Learning Disability Quarterly*, *34*(4), 262–272. https://doi.org/10.1177/0731948711421762

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Pennequin, V., Sorel, O., Nanty, I., & Fontaine, R. (2010). Metacognition and low achievement in mathematics: The effect of training in the use of metacognitive skills to solve mathematical word problems. *Thinking & Reasoning*, *16*(3), 198–220. https://doi.org/10.1080/13546783.2010.509052

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, *41*(4), 219–225. https://doi.org/10.1207/s15430421tip4104_3

Quigley, A., Muijs, D., & Stringer, E. (2018). Metacognition and Self-regulated Learning: Guidance Report. In *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit*.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, *26*(1–2), 113–125. https://doi.org/10.1007/978-94-017-2243-8_1

Schraw, G., & Dennison, R. S. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

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

Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning*, *19*(1), 46–55. https://doi.org/10.1046/j.0266-4909.2003.00005.x

Veenman, M. V. J. (2015). Metacognition. In P. Afflerbach (Ed.), *Handbook of Individual Differences in Reading: Reader, Text, and Context*. Routledge. https://doi.org/10.4324/9780203075562.ch3

Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences*, *15*(2), 159–176. https://doi.org/10.1016/j.lindif.2004.12.001

Wang, A. Y., Fuchs, L. S., Fuchs, D., Gilbert, J. K., Krowka, S., & Abramson, R. (2019). Embedding Self-Regulation Instruction Within Fractions Intervention for Third Graders With Mathematics Difficulties. *Journal of Learning Disabilities*, *52*(4), 337–348. https://doi.org/10.1177/0022219419851750

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What Influences Learning? A Content Analysis of Review Literature. *Journal of Educational Research*, *84*(1), 30–43. https://doi.org/10.1080/00220671.1990.10885988

Wilson, D. B. (n.d.). *Campbell Collaboration Effect Size Calculator*. Retrieved January 17, 2020, from http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD3.php

Zhu, N. (2015). Cognitive strategy instruction for mathematical word problem-solving of students with mathematics disabilities in China. *International Journal of Disability, Development and Education*, *62*(6), 608–627. https://doi.org/10.1080/1034912X.2015.1077935