Mathematics Education Research Journal (2023) 35:821–848
https://doi.org/10.1007/s13394-022-00414-0

ORIGINAL ARTICLE

Successful teaching practices for english language learners in multilingual mathematics classrooms: a meta-analysis

Shweta Sharma[1] · **Sashi Sharma**[2]

Received: 25 October 2021 / Revised: 23 February 2022 / Accepted: 6 March 2022 /

Abstract

Due to rapid immigration, many children worldwide are learning mathematics in a second or additional language. This language diversity can be challenging for both teachers and students and carries profound implications for mathematics educators. Research shows that teachers use various ways to support English Language Learners. Research on multilingualism in mathematics classrooms has often focused on qualitative research. This meta-analysis aims to explore the statistically effective successful teaching practices from the studies using quantitative or mixed-method research approaches and aims to inform the research field in a cumulative manner. The specific research question that guided this meta-analysis is: What is the evidence regarding successful teaching of mathematics for Year 1–10 English Language Learners from 2009–2019 in countries where curricula are delivered predominantly in English? Four successful intervention categories were identified: Dual Language Programmes, Curriculum integration, Teacher Professional Development, and Cognitively Focused Interventions. The paper concludes with recommendations for practice and further research in this area.

Keywords School mathematics · English language learners · Meta-analysis · Successful mathematical practices

Introduction

Multilingual classrooms are an increasing feature of schools worldwide, in part due to immigration as a response to poverty and war but also as part of efforts to maintain minority or indigenous home country languages (Education Review Office, 2018;

✉ Shweta Sharma
shweta.sharma@canterbury.ac.nz

Sashi Sharma
sashi@waikato.ac.nz

[1] University of Canterbury, Christchurch, New Zealand

[2] University of Waikato, Hamilton, New Zealand

European Commission, 2015). Multilingualism implies covert or overt presence of two or more languages in a classroom (Barwell et al., 2016). In multilingual classrooms, students may speak one language at home and another language at school. Additionally, teachers and students may not share a common language or cultural background. Alternatively, there may be multilingual classrooms where some or all the students may be learning the language of instruction as a second language. Aotearoa New Zealand, for example, is a superdiverse bicultural nation with diverse ethnicities of its people and languages spoken (Education Review Office, 2018). New Zealand classrooms are places where learners bring their different linguistic, cultural, and everyday experiential backgrounds. This language diversity can be challenging for both teachers and students and carries profound implications for mathematics educators. Importantly, full participation of English Language Learners (ELLs) into the learning community is essential, not only with regard to issues of equity, but also to recognise the assets ELLs bring to a classroom environment (Lesser et al., 2016).

English Language Learners are students who come from non-English-speaking backgrounds, who are not proficient in English, and who require specialised or modified language instruction. These learners are known by a variety of names in the education community: English Language Learners, English Speakers of Other Languages (ESOL) learners, English as foreign language speakers, English as second language learner, English Learners (or ELs), Limited English Proficient (LEP) students, non-native English speakers, language-minority students, and/or bilingual students or emerging bilingual students. We use the term English Language Learners (ELLs), in this meta-analysis, because of its alignment with the majority of literature reviewed for this review. However, we do realise that the phrase can be problematic. For example, it identifies students by what they lack rather than focusing on the funds of resources that students bring to the school. We need to highlight students' language as a resource rather than language as problematic. We also note that the term "English Language Learners" is preferable to older terms such as "Limited English-proficient" which have been used in research and policy documents.

Thus, in this paper, we identify successful teaching practices for Year 1–10 English Language Learners (ELL) from 2009 to 2019 in countries where curricula are delivered predominantly in English.

Background

Mathematician Halmos (1980) explains, "that the mathematician's main reason for existence is to solve problems, and that, therefore, what mathematics really consists of are problems and solutions" (p. 519). Mathematics fosters critical thinking and problem-solving. Therefore, it is important for success in school and in society, not just for some students but for each and every student, as argued by Moses and Cobb (2001):

> It's not cool or hip to be completely illiterate in math. The older generation may be able to get away with it, but the younger generation coming up now can't – not if they're going to function in society, have economic viability, be in a position to

> meaningfully participate, and have some say-so in decision making that affects their lives (p. 14).

Educators struggle to meet the needs of an increasingly diverse student population. For example, New Zealand's Ministry of Education regards students' lack of numeracy and literacy skills as major impediments to sustained personal, and future national economic growth (Ministry of Education, 2014). Despite government policies of reviewing and adapting successful educational systems, mathematics results have been disappointing. Trends in International Mathematics and Science Study (TIMSS) outcomes plateaued in 2016 (Mullis et al., 2016). The Year 9 students' scores in the TIMSS fell by the largest margins since the study began in 1994. Their mathematics score fell 11 points to 482 where 500 is the midpoint (RNZ, 2020). The Programme for International Student Assessment (PISA) 2018 results for New Zealand students are similar to PISA 2015 and PISA 2012 results (Hipkins, 2019). However, there has been a decline in mathematics performance since 2003 (Ministry of Education, 2019). Hence, it is imperative that teachers understand what effective mathematics teaching looks like and what teachers can do to break this pattern.

Moreover, there have been significant shifts in the way teaching and learning of mathematics is conceptualised internationally. The American policy document, Principles and Standards for School Mathematics states that if students are to learn to "construct mathematical arguments and respond to others arguments", then creating an environment that fosters these kinds of activities is essential (National Council of Teachers of Mathematics, 2000, p. 18). This reform aligns with the vision promoted by The Partnership for 21st Century Skills (2009) which argues that educational outcomes for students in the twenty-first century must focus on communication and collaboration across multiple languages and cultures. However, learning and communicating about mathematics is not an easy task for many students. For instance, students may fail to understand the content in textbooks or may fail to understand the instruction for assessment. The difficulty to comprehend mathematics language may further compound or aggregate for learners who need to overcome language barriers (Barwell, 2020; Clarkson, 2007). It must be noted that many children around the world learn mathematics in a second language.

Learning mathematics requires learning its associated language. Learning the language used for defining mathematical concepts and communicating mathematical ideas plays a key role in mathematics learning and teaching (Anthony & Walshaw, 2007). Mastering a new mathematics topic requires understanding and appropriately using the language of the discipline through listening, reading, writing, and speaking. Responses to multilingualism in mathematics education research vary. Planas and Setati-Phakeng (2014) describe three perspectives that impact on the development of language policies and multilingual classroom practices: language-as-problem, language-as-right, and language-as-resource. The language-as-problem perspective considers language as something that creates challenges that need to be resolved. In this view, teachers may view students' limited English as a limitation to be overcome through a focus on intensive language teaching. Language-as-right emphasises the protection of minority language groups since everyone has the right to be educated in her/his home language. Planas and Setati-Phakeng (2014) further note that while language-as-right supports the use of the students' home languages as the language of learning and instruction, this initiative is often paired with the stigma of the home language being a “non-English” language. The pedagogical strategies and policies based on language-as-problem and language-as-right can have unintended effects on different language groups of students by decreasing their access to classroom learning opportunities and interaction. By contrast, the language-as-resource perspective addresses both the stigma and the problem through actively encouraging the use of multiple languages during mathematics teaching. Planas and Setati-Phakeng (2014) see a language-as-resource approach as increasing the learning opportunities of all learners by focusing on both mathematics and language as being connected in the teaching and learning process.

Grappling with how language is used in mathematics can present challenges for any student (Kazima, 2006; Lesser & Windsor, 2009; Warren & Miller, 2015). However, ELLs in English-medium classrooms face additional challenges (Barwell, 2009; Kazima, 2006; Saxe & Sussman, 2019) because they need to simultaneously learn everyday English and mathematical English and to differentiate between the two types of English language use (de Oliveria & Cheng, 2011; Schleppegrell, 2011). Supporting ELLs can pose challenges for teachers all the more so because mathematics classes are often multilingual in nature. In a multilingual setting, students may miss out on learning because they may be spending too much time shifting between informal and formal ways of communicating ideas while trying to understand the instructions and questions. In addition to this, ELLs are often marginalised in mathematics classrooms due to these language challenges (de Araujo et al., 2018). Many empirical studies have shown that activating the first languages of ELLs helps to give access to mathematics (Barwell, 2009; Moschkovich, 2018; Planas & Setati-Phakeng, 2014). For example, Clarkson (2007) explains how English Language Learners may comprehend target language texts using their first learnt language (L1). He claims that the first language scaffolds semantic processing, whereas if a learner were to process the input exclusively in second language/formal language of instruction, then s/he might run into difficulties understanding syntactically complex sentences.

Many studies on effective teaching and learning of mathematics have been conducted worldwide, but many of these studies are small scale and qualitative. Following an extensive review of research on ELLs in K–12 mathematics, de Araujo et al. (2018) suggested that review studies need to be conducted that go beyond individual case studies and inform a work of a more cumulative nature. This idea was echoed by Schmidt (1992), “Many discoveries and advances in cumulative knowledge are being made not by those who do primary research studies, but by those who use meta-analysis to discover the latent meaning of existing research literature” (p. 1179).

This meta-analysis aims to provide an overview of quantitative and mixed-method peer-reviewed studies published from 2009 to 2019 as a guide for future practice and research. This paper provides an evidence-based review of studies to inform the effective teaching and learning approaches in mainstream primary and secondary mathematics education in countries where curricula are delivered predominantly in English. The following research question guided this evidence review:

> What is the evidence regarding successful teaching of mathematics for Year 1–10 English Language Learners in countries where curricula are delivered predominantly in English?

The meta-analysis process

Meta-analysis is essentially a systematic literature review. However, in addition to the narrative summary that is conducted in a systematic review, analysts conducting meta-analysis numerically pool the results of similar studies to arrive at a summary estimate or effect size (Reljić et al., 2015). The purpose of meta-analysis is to address a broader question, thus, enabling the generalisability of findings across similar studies (Borenstein et al., 2009). As Ellis (2010) explains it, meta-analysis is a set of procedures for systematically reviewing research, examining a particular effect, and combining the results of independent studies to estimate the size of the effect of treatment in the population. That is, meta-analysis is a quantitative procedure that is used to statistically combine the results from studies on the same topic and allows us to reach some more statistically significant conclusions regarding the effects or outcomes of a given treatment, project, or programme (Cooper, 2009).

There are two models that are used to conduct meta-analysis—the fixed-effect model and the random-effect model (Ellis, 2010). Fixed-effect models are used to conduct meta-analysis of the studies with quite similar findings. In this case, it is assumed that the variation among the results of studies is due to sampling error. In other words, the fixed-effect model assumes that there is one true effect size that underlies all the studies in the analysis, and that all differences in observed effects are due to sampling error. However, studies that are uniform on account of intervention used, selection of the population, and the outcome often report varied results. In this case, a random-effect model is used (Borenstein et al., 2009). The random-effect model allows the true effect sizes to differ. Moreover, it enables the explanation of the variances caused in the effects of treatment across studies. For example, the effect size might be higher (or lower) in studies where the participants are older, or more educated, or when a more intensive variant of an intervention is used (Byun & Joung, 2018). In this model, the effect sizes in the studies that actually were performed are assumed to represent a random sample from a particular distribution of these effect sizes. The selection of the model is critically important, and it affects the computations, and the analysis and the interpretation of the statistics.

Since the studies reported in this meta-analysis are heterogeneous in their reporting of findings, a random-effect model has been used. In reporting the results of the present meta-analysis, forest plots for the overall combined effect and combined effect for different teaching intervention categories are presented. Forest plots visually present the effect size and confidence interval for each study, the weights assigned to each effect size, and the estimate of summary effect (Borenstein et al., 2009). Effect size is a value that "reflects the magnitude of the treatment effect or the strength of a relationship between two variables" (Borenstein et al., 2009, p. 3). In this report, hedge's *g* is reported as the appropriate effect size (Reljić et al., 2015) with lower and upper limits within square brackets, i.e. []. As mentioned earlier, hedge's g is presented as the index of effect size, and its value of as 0.15, 0.40, and 0.70 should be interpreted as small, medium, and large effects, respectively (Lovakov & Agadullina, 2021).

In a forest plot (as shown in Fig. 1), the average effect size of each individual study is depicted with a square box. The size of the square box indicating the effect

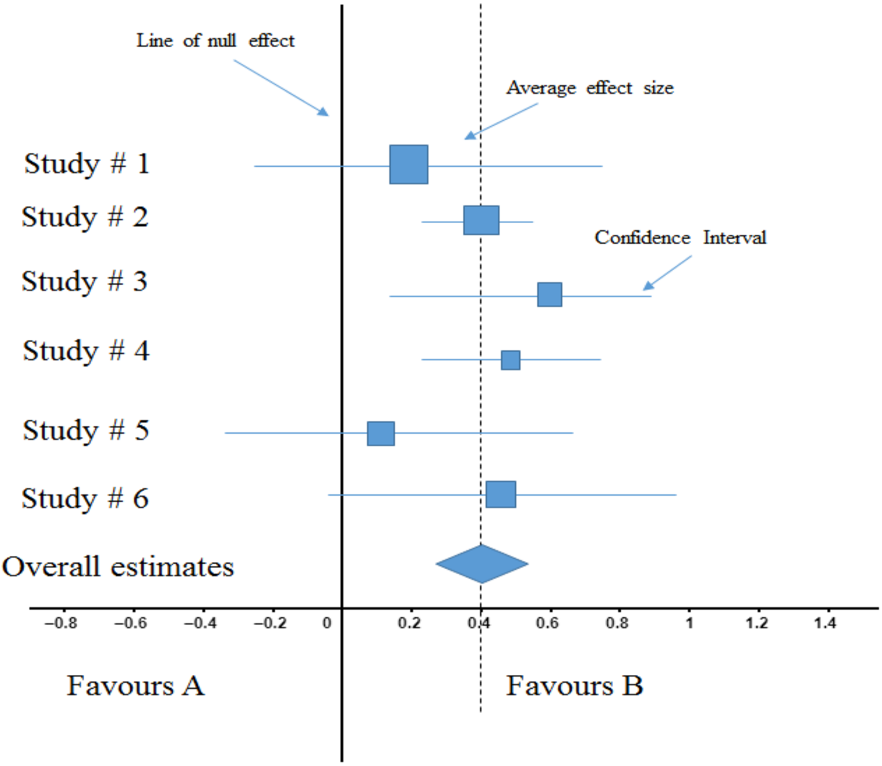


Fig. 1 Example of a forest plot

size is directly proportional to the weight assumed by the study in the meta-analysis. Study weights depict how each study varies from others in addition to the effect estimate of the study. That is to say, if the variance of a study is high, the size of the square would be small. Moreover, each horizontal line passing through the square box indicates the confidence interval. The confidence interval reflects the precision with which the effect size is calculated for each study. The 0.00 vertical axis displays the line of null effect. The overall effect size or a summary estimate is depicted using a diamond shape. The position of the overall effect size (in the shape of diamond) shows the direction and magnitude of treatment effect. The diamond does not have a line that corresponds to the 95% confidence interval; instead, the width of the diamond represents the 95% confidence interval band around it. Thus, a forest plot visually presents effect estimates of individual studies distributed around a null value and the overall effect estimates.

Meta-analysis proceeds through a series of steps, typically, (i) searching and identification, (ii) screening and selecting, (iii) coding and eligibility, and (iv) conducting meta-analysis. These steps are detailed below. Each of the steps were followed to answer the research question. Flow charts of the process, which follow the PRISMA recommendations (Moher et al., 2009), is also presented.

Step 1: Searching and identification of studies

To ensure rigour, a comprehensive search strategy was devised utilising systematic techniques such as those outlined in PRISMA methodology (Moher et al., 2009) to identify the relevant studies. For the research question, a logic grid was used to identify the appropriate key search terms and synonyms. Boolean logic (AND, OR, NOT) was used to develop a combination of keywords to be used for searching the selected databases. These databases included: Education Database (ProQuest), Education Source (EBSCO), ERIC (Educational Resources Information Center), Scopus, JSTOR, and ScienceDirect. In addition, we also searched using Crossref.org (Crossref is an official Digital Object Identifier Registration Agency of the International DOI Foundation) with the same set of keywords to identify additional sources for meta-analysis and to increase reliability of search findings. To screen the potentially relevant studies at this step, the following inclusion and exclusion criteria were employed. The inclusion criteria were:

- Peer-reviewed journal articles and conference proceedings.
- Participants from Years 1–10.
- Quantitative and mixed-method studies.
- Articles pertaining to English language learners and speakers of other languages.
- Studies with participants from these eight countries: Canada, USA, Australia, Scotland, Ireland, England, Wales, and New Zealand.

The following exclusion criteria were applied:

- All book chapters, teaching activity articles, literature reviews, and theoretical papers.
- Articles in language apart from English.

Step 2: Screening and selecting studies

At this step, studies identified at step 1 were further screened by reading abstracts to select the potentially relevant studies. The search for articles relevant to this question focused on English Language Learners. Search terms were: (mathematics OR algebra OR statistics OR calculus OR geometry OR trigonometry OR arithmetic OR "number system") AND (teach*) AND ("English language" OR "English language" OR ESL OR ESOL) AND (quantitative OR "mixed methods" OR "mixed-methods"). Initially 1847 articles appeared in the search, after title and keywords screening and removing duplicate records 467 articles were identified. Screening of abstracts reduced this number to 210. Further screening identified 61 articles for full-text review. Finally, the full-text review of 61 articles found 07 studies meeting the criteria of inclusion for meta-analysis (as shown in Fig. 2).

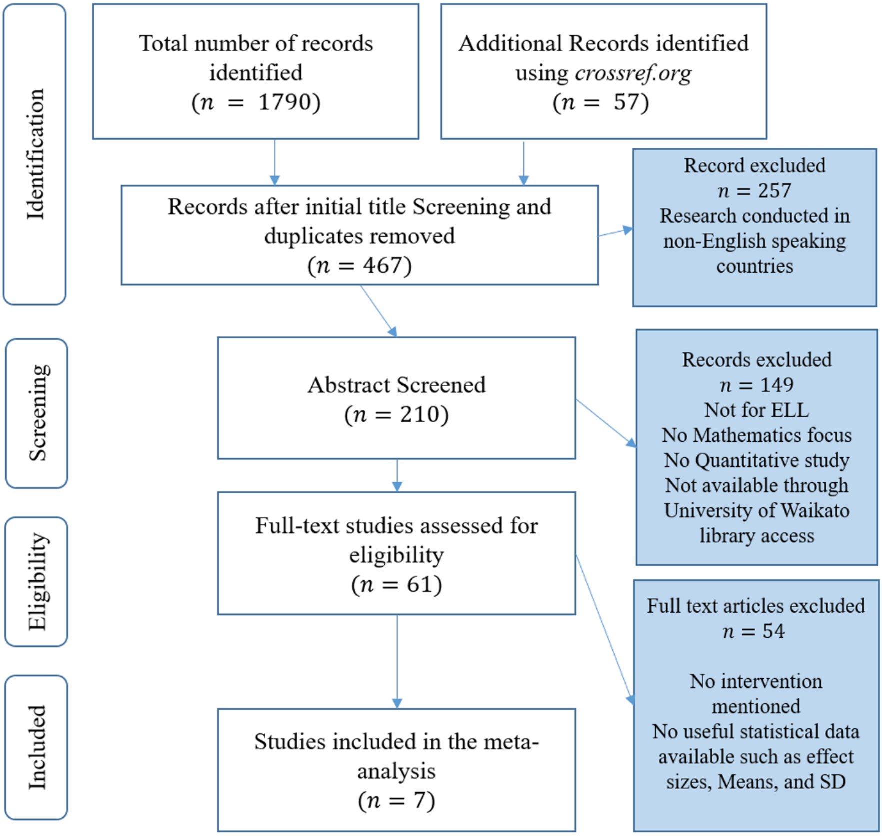


Fig. 2 Screening of Research Studies

Step 3: Coding of the selected studies for analysis

For the purpose of coding, a coding scheme was developed. Each study was coded as per the coding scheme in Table 1. A small sample of the selected studies were coded independently by two researchers to gauge inter-rater reliability. This process was central in reaching agreement as to whether a study was relevant for meta-analysis. During this process, the research team achieved 90% of inter-rater reliability (Borenstein et al., 2009). Once inter-rater reliability was achieved, both the authors coded all the selected studies using the coding scheme outlined in Table 1. A similar process was used when identifying the intervention categories. Two members of the research team identified the categories independently, and then, a discussion was held to come to a consensus.

Step 4: Conducting meta-analysis of the selected studies

Comprehensive Meta-analysis (CMA) software was used to conduct a meta-analysis of the selected studies. To calculate the combined effect of all the teaching practices

Table 1 Description of codes used for included studies

Codes	Description
Initials	Initials of the person screening the study
Study	Name of the author and the title of the study
Database	Database from where the study is taken
Document type	Whether the document is an article or conference paper
Purpose/ Aim of the study	What was the main purpose of the research?
Topic Studied (in the study)	What topics were the focus of the study? E.g., problem-solving, mathematical reasoning, Algebra, Number system or any other
Type of Study	General study design. Possible categories: (1) one-group pre–post-measurement, (2) pre–post-comparison group design, (3) comparison group design without pre-measurement
Intervention	What intervention was used? What teaching strategy was used to conduct the study?
Participants (Year Level)	Participants' year level. Possible categories: (1) primary/elementary (kindergarten—grade 6), (2) secondary/high school (grades 7–12), (3) mixed (kindergarten–grade 10)
Sample Size (n)	Sample size of the study students/teachers
Data collection tools	Instruments used for data collection. Possible categories: (1) survey, (2) interview, (3) observation, (4) assignment/test
Statistical Values (M, SD, r)	Outcome data applicable for meta-analysis in the form of Means, Standard deviation and correlations
Statistical Measure	A summary (means, mode, total, index, etc.) of the individual quantitative variable values for the statistical units in a specific group (study domains)
Effect Sizes	Measures the strength of the relationship between two variables on a numeric scale
Reliability of data collection tool	Reliable measures/data collection process (inter-observer agreement of ICC/Kappa $\geq$ .70). Possible categories: (1) not included (i.e. reliability lower than cut-off values or not reported), (2) included in study

identified in the selected studies, a meta-analysis of all selected studies was performed assuming both fixed-effect and random-effect models. The test for heterogeneity informed us that the selected studies were heterogeneous in nature. The variability of findings can be attributed to the different methods of data collection, presence of various moderator variables, for example, age, gender, Year-level, and different interventions among other factors. Hence, the random-effect model was used to calculate a mean effect size. Seven studies in total were selected for this meta-analysis; however, six of these studies are from USA only one study from Australia; hence, subgroup analysis could not be conducted (Hak et al., 2016).

An overview of the seven selected studies for the research question (author(s); publication year; research design, context of study, sample size, and intervention or teaching strategy) is presented in Table 2.

Next, based on the coding of studies, the studies were clustered into different intervention categories and studies pertaining to one kind of intervention were

Table 2 Overview of selected studies in alphabetical order ($n=7$)

Author (Year)	Database	Research Design	Context of Study			Sample Size	Intervention/ Teaching Strategy
			Country	Education level	Focus of Study		
Anderson et al. (2018)	ProQuest	Mixed Methods	USA	Grade 5	Impact of change in teacher's mathematics mindset on students' mathematics performance	Teacher (n) = 40, Students (n) = 1068	Blended Mathematical mindset professional development approach
Clements et al. (2013)	JSTOR	Quantitative	USA	Grade 1	Overall mathematics achievement	N = 1079 from 42 schools	Technology -enhanced, Research-based, Instruction, Assessment, and Professional Development (TRIAD) model
Jitendra et al. (2013)	JSTOR	Quantitative	USA	Grade 3	Word problem-solving, Number computation sense	Students N = 136, Teacher n = 20	Schema-based Intervention
Matthews and López (2019)	Scopus	Mixed Methods	USA	Grade 3 to 5	Mathematics Assessment	Students n = 568, Teacher n = 33,	Asset-Based Pedagogy
Saxe and Sussman (2019)	EBSCO	Quantitative	USA	Grade 4 and 5	Students' mathematics assessment	N = 571	Learning Mathematics through Representations (LMR) curriculum
Vela et al. (2017)	ProQuest	Quantitative	USA	Grade 3	Mathematics assessment	N = 2279	Dual language two-way immersion programme
Warren and Miller (2015)	ProQuest	Quantitative	Australia	Year 1 and 2	Students' mathematics achievement	461	Representations, Oral Language and Engagement in Mathematics (RoleM) Learning activities

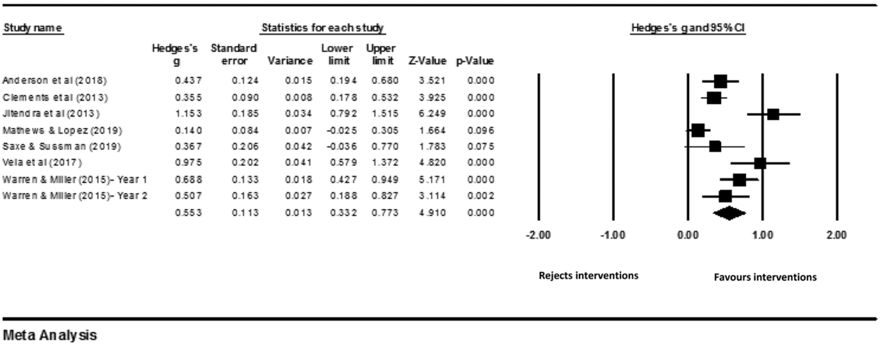
subjected to independent meta-analysis using a random-effect model. Therefore, an overall effect size, individual effect size for each intervention category along with effect sizes for individual intervention/study are reported. The next section presents the results of the meta-analysis. Based on the identified interventions, a set of focused recommendations is provided for the development and support of teacher knowledge and practice.

Overall results

Seven studies were subjected to meta-analysis for identifying successful teaching practices for English Language Learners. Figure 3 presents the effect size for individual studies along with a combined random-effect size for all studies as a forest plot. One study, Warren and Miller (2015), was coded for two different year levels. Warren and Miller (2015) provided the effect sizes for Year 1 and Year 2 students.

The summary effect size shows that, overall, the interventions in the selected research studies have a moderate yet significant positive effect on English language students' mathematics academic performance with hedge's g as 0.553 with 95% CI [0.332, 0.773].

The studies were sub-grouped into four intervention categories. The four interventions are (i) Dual Language Programmes, (ii) Professional Development for Teachers, (iii) Curriculum Intervention, and (iv) Cognitively Focused Interventions. Table 3 shows the identified categories and the studies pertaining to each intervention. The effect sizes for each intervention category are discussed separately in brief in the following sections.



Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderson et al (2018)	0.437	0.124	0.015	0.194	0.680	3.521	0.000
Clements et al (2013)	0.355	0.090	0.008	0.178	0.532	3.925	0.000
Jitendra et al (2013)	1.153	0.185	0.034	0.792	1.515	6.249	0.000
Mathews & Lopez (2019)	0.140	0.084	0.007	-0.025	0.305	1.664	0.096
Saxe & Sussman (2019)	0.367	0.206	0.042	-0.036	0.770	1.783	0.075
Vela et al (2017)	0.975	0.202	0.041	0.579	1.372	4.820	0.000
Warren & Miller (2015)- Year 1	0.688	0.133	0.018	0.427	0.949	5.171	0.000
Warren & Miller (2015)- Year 2	0.507	0.163	0.027	0.188	0.827	3.114	0.002
	0.553	0.113	0.013	0.332	0.773	4.910	0.000

Fig. 3 Overall Meta-Analysis

Table 3 Intervention categories with relevant studies

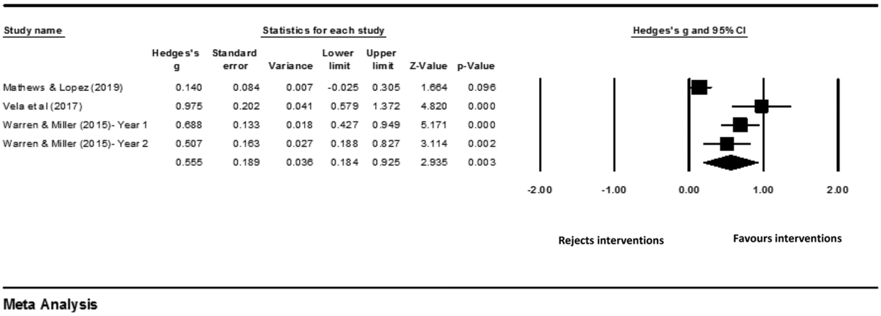
Intervention categories	Number of studies	Studies
Dual Language Programmes	3	Matthews and López (2019) Vela et al. (2017) Warren and Miller (2015)
Professional Development for Teachers	2	Anderson et al. (2018) Clements et al. (2013)
Curriculum Intervention	1	Saxe and Sussman (2019)
Cognitively Focused Interventions	1	Jitendra et al. (2013)

Intervention category 1: Dual language programmes

The intervention theme of "Dual Language Programmes" was the most prominent in the selected studies, being mentioned in three of the seven studies. Dual language education, often called bilingual education, refers to academic programmes that are taught in two languages (García & Leiva, 2014; Reljić et al., 2015). The proposal is that Dual Language Programmes foster bilingualism, biliteracy, enhanced awareness of linguistic and cultural diversity, and high levels of academic achievement (Planas & Setati-Phakeng, 2014). The three studies in the intervention category are: Matthews and López (2019), Vela et al. (2017), and Warren and Miller (2015). Figure 4 shows the individual and combined effect size of these studies from a random-effect model. The Warren and Miller study appears twice because they reported changes for Year group 1 and Year 2 students. The combined effect size is 0.555 with 95% CI [0. 184, 0.925], which is moderate in magnitude.

The following teaching practices are included in this intervention category:

- Asset-based pedagogy (Matthews & López, 2019).
- Dual language two-way immersion programme (Vela et al., 2017).
- Representations, Oral Language and Engagement in Mathematics Learning activities (Warren & Miller, 2015).



Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Mathews & Lopez (2019)	0.140	0.084	0.007	-0.025	0.305	1.664	0.096
Vela et al (2017)	0.975	0.202	0.041	0.579	1.372	4.820	0.000
Warren & Miller (2015)- Year 1	0.688	0.133	0.018	0.427	0.949	5.171	0.000
Warren & Miller (2015)- Year 2	0.507	0.163	0.027	0.188	0.827	3.114	0.002
	0.555	0.189	0.036	0.184	0.925	2.935	0.003

Fig. 4 Meta-Analysis of the Studies with Dual Language Programme

The first intervention within this category is that of asset-based pedagogy (ABP) (Matthews & López, 2019). The individual effect size for this intervention is 0.140 (small) with 95% CI [-0.025, 0.305]. This pedagogical approach focuses on *cultural content integration* (CCI) as well as respecting students' *heritage language*. The aspect of cultural content integration allows incorporation of students' cultural content into classroom instruction. This integration particularly aims to bring marginalised cultural content within mainstream classroom learning. The second component of bringing heritage language (in this case, Spanish) within the classroom aims to develop positive student identities. Using explanatory sequential mixed-method design, Matthews and López (2019) examined whether teachers' use of Spanish language during instruction mediates the relation between CCI and grade 3–5 students' growth in mathematics achievement. Mathematics achievement assessments aligned with state's academic standards. Significant impact of ABP on grade 3–5 students' mathematics achievement was found.

The second intervention involves a dual language programme as described in Vela et al. (2017). The purpose of the Vela et al. (2017) study was to determine whether there is a difference among ELLs who were enrolled in one of three programmes: (1) a transitional bilingual programme, (2) a dual language two-way immersion programme, or (3) a regular programme with immersion into all-English instruction in reading and math. In total, 2279 grade 3 students from an elementary urban school participated in this study. The researchers analysed and compared the State of Texas Assessments of Academic Readiness (STAAR) mathematics standardised tests used in public primary and secondary schools to assess a student's achievements and knowledge learned in the grade level. The dual language two-way immersion group scored significantly higher than the all-English immersion group and the transitional instructional group. The structure of Dual Language Programmes varies, but they all provide at least 50% of instruction in the partner language at all grade levels beginning in pre-K, Kindergarten, or first grade and lasting through at least five years (Tran et al., 2015). The individual effect size of dual language two-way immersion intervention is 0.975 (i.e. large in magnitude) with 95% CI [0.579, 1.372]. The two-way immersion programme involved an instructional programme in which students are taught in two languages simultaneously.

Representations, Oral Language and Engagement in Mathematics (RoleM) programme was identified as the third intervention in this category. Warren and Miller (2015) studied the impact of purposefully developed learning activities on students' mathematics outcomes in the first three years of formal schooling (Foundation, Year 1 and Year 2). The RoleM learning activities are based on a socio-constructivist perspective. The aim is for conceptually oriented, culturally appropriate as well as cognitively demanding activities that support students' learning pathways. Warren and Miller (2015) state that the RoleM learning activities take account of:

- *Learning pathways*——providing a gradual progression along a learning path, with the teacher first modelling what is required, followed by children of similar ability working in groups and finally children working on an individual basis.

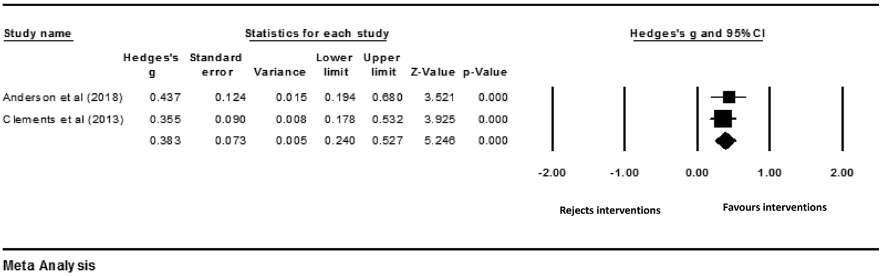
- *Integrated experiences*——involving listening, reading, writing, recording, manipulating, physically moving and speaking about the concepts to enhance children's transference of skills.
- *Multirepresentations*——using and linking concepts to a variety of mathematical representations including number lines, charts, concrete and symbolic.
- *Language building*——encouraging children to move between home language, mathematical language and SAE as they communicate their mathematical learning.
- *Engaging and focused*——ensuring that the materials were visually stimulating while specifically focused on the mathematical concept under consideration.
- *Making connections*——linking resources to other mathematical concepts and with children's home and community environment (Frigo & Simpson, 2001; Jackson & Cobb, 2010; Warren & deVries, 2009; as cited in Warren & Miller, 2015, p. 197).

The participating group ($n=461$) comprised 328 English language learners (132 Foundation year, 119 Year 1, and 77 Year 2 students) and 133 mainstream pupils. To ascertain the impact of the learning experiences, pre- and post-tests were conducted at the commencement and completion of each school year. The results indicate that all of the children significantly improved with English as a second language pupils showing the greatest gains and achieving norm-referenced expectations for their age. The individual effect sizes of this intervention for Year 1 and Year 2 are 0.688 and 0.507 with 95% CI with upper and lower limits of [0.427, 0.949] and [0.188, 0.827], respectively. These effect sizes are moderate in statistical power.

Intervention category 2: Professional development for teachers

For the "Professional Development for Teachers" category, the two selected studies are Clements et al. (2013) and Anderson et al. (2018). Figure 5 shows the individual and combined effect size of these two studies.

The combined effect size of the Professional Development for Teachers intervention category is moderate effect size with hedge's g as 0.383 (as this value is closer to 0.4) with 95% CI [0.240, 0.527] for professional development for teachers as an



Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Anderson et al (2018)	0.437	0.124	0.015	0.194	0.680	3.521	0.000
Clements et al (2013)	0.355	0.090	0.008	0.178	0.532	3.925	0.000
	0.383	0.073	0.005	0.240	0.527	5.246	0.000

Fig. 5 Meta-Analysis of the Studies with Professional Development Interventions

intervention category. Although the effect size is moderate in magnitude, the two-professional development programmes mentioned in the studies signal the value of Teacher Professional Development in helping promote the mathematics achievement of ELL students.

The teaching practices within this intervention category are:

- Blended mathematical mindset professional development approach (Anderson et al., 2018).
- Technology-enhanced, Research-based, Instruction, Assessment, and professional Development (TRIAD) model (Clements et al., 2013).

The first intervention in this category is the Blended mathematical mindset professional development approach reported in the study by Anderson et al. (2018). The effect size for this study is 0.437 (moderate in magnitude) with 95% CI [0.194, 0.680]. Anderson et al. (2018) investigated the impact of a blended mathematical mindset professional development approach on teachers' beliefs about being "math person", or not. The researchers made use of research on mathematics education and mindset and also from neuroscience. Teachers participated in online and in-person professional development. The overall aim of the professional development was to challenge stereotypical beliefs and myths about there being a "math person" and to focus on the development of a "growth mindset" (Dweck, 2006) for teachers themselves and in relation to their work with their students. Teachers were provided with opportunities to explore how beliefs and myths can act as hurdles in mathematics teaching and learning. The impact of this professional development was assessed using a control–experimental research design as part of a mixed-method study. Anderson et al. (2018) identified a statistically significant positive shift in student beliefs, teachers' instructional practices, and students' achievement state mathematics tests. They attributed the success of the intervention to the use of different forms of professional development focused on challenging the myths about learning held by teachers and learners and teachers having space for identity work to do with themselves as mathematical learners.

The second intervention in this category is the Technology-enhanced, Research-based, Instruction, Assessment, and Professional Development (TRIAD) model (See Sarama et al., 2012), as described in Clements et al. (2013). The TRIAD model is based on a network of influences theory (Sarama et al., 1998), and it is aimed at developing and scaling up a research-based curriculum. The TRIAD model has four major components: professional development for teachers, research-based learning trajectories, research-based curriculum and instructional strategies, and assessment. Sarama et al. (2012) argued that the professional development component of the TRIAD model enables teachers to become responsive and to develop their understanding of teaching and learning processes along with curriculum and assessment by informing them about research-based models of' student mathematical thinking and learning. Moreover, through TRIAD professional development, teachers learn about students' learning trajectories as marked by a goal, and a developmental progression of thinking and instruction aimed at promoting students' movement along that learning trajectory. The instruction is focused on the teaching strategies in accordance with the curriculum that promotes active mathematical learning of all students including ELLs. Clements et al. (2013) found that the TRIAD model was associated with significantly improved mathematics achievement of ELLs who had limited English language proficiency. The individual effect size of this intervention is 0.355 with 95% CI [0.178, 0.532]. The effect size is moderate in magnitude.

Intervention category 3: Curriculum intervention

Curriculum Intervention category involves the intervention where a new curriculum approach is undertaken. The strategy identified within this intervention category is the *Learning Mathematics through Representations* (LMR) curriculum approach, as explained by Saxe and Sussman (2019). The LMR curriculum consists of 19 lessons on integers and fractions. In this Curriculum Intervention, the number line is used as the primary representational context. To develop an understanding of integers, students engage with activities which require them to create, define, and reflect on numbers using the units and multiunits to the right of zero on the number line. Similarly, for the negative integers, students are required to reflect on the units and multiunits to the left of the number line. The students move to fractions from integers as the lessons proceed. Each LMR lesson involves teachers gauging students' thinking and building upon this in their instructional practice. Saxe and Sussman (2019) analysed the efficacy of LMR on the mathematics learning of ELLs as well as English proficient speakers, 571 grade 4 and 5 students in all. Four specialised assessments were administered in the months of September, October, December, and May. They found that ELLs' performance improved steadily in LMR classes. The individual effect size for this intervention is 0.367 with 95% CI [-0.036, 0.770]. This can be interpreted as moderate effect size as the value is closer to 0.4. However, Saxe and Sussman (2019) argue that use of representations increases the likelihood that all students, including ELLs who may have difficulty accessing traditional mathematics curricula, will have multiple opportunities over time to engage with complex mathematical ideas and build proficiency.

Intervention category 4: Cognitively focused interventions

Cognitively Focused Interventions is concerned with the teaching practices that aim to improve learning processes including attention, memory and metacognition (Davenport et al., 2019). Therefore, the strategies identified in this intervention category work on the principles of how the knowledge is received, processed, retained and used. Only one study is identified in this category: Jitendra et al. (2013). The identified intervention is Schema-based Instruction (Jitendra et al., 2013). Jitenda et al. (2013) assert that as the "multiple elements of information are grouped into and conceptualised as a single schema, recognizing a problem's schema reduces the working memory load during cognitive processing" (p. 22). The researchers investigated the impact of small-group tutoring on the mathematical problem-solving and achievement of grade 3 students. In the study, the small-group tutoring was developed on the principles of schema theory. The study was conducted in 12 elementary schools in a large urban district. Thus, the intervention used is referred to as a Schema-based instruction (SBI). Jitendra et al. found that SBI positively impacted students' performance on word problem-solving and number operations, and their overall achievement in mathematics assessment. SBI emphasises the acquisition and identification of the underlying structure of problems along with the sub-components of information provided in a word problem as central to successful problem-solving. Jitendra et al. (2013) asked students to "think aloud" while solving a problem so that the behaviours of good problem-solvers could be explored as they engaged in the process of problem comprehension and problem solution. Employed this way, the use of SBI enabled the teachers to gauge students' metacognitive strategies including organisation of data, planning of solutions, execution of plans, and checking results (Coldberg & Bush, 2003) in addition to algorithms and heuristics. The individual effect size for this intervention is large with hedge's g as 1.153 with 95% CI [0.792, 1.515].

The findings presented four intervention categories with seven successful teaching practices that were identified in this meta-analysis for improving learning outcomes for ELLs. Discussion of these findings is presented now.

Discussion

The purpose of this meta-analysis was to analyse quantitative and mixed-method studies of effective teaching and learning approaches in primary and secondary mathematics education in countries where curricula are delivered predominantly in English. The following research question guided this evidence review:

> What is the evidence regarding successful teaching of mathematics for English Language Learners and speakers of other languages in countries where curricula are delivered predominantly in English?

The meta-analysis located and selected seven studies that met the inclusion criteria, with these seven studies providing eight effect sizes. The sample sizes in the reported studies ranged from three to 2270. Interestingly, all studies included participants from primary schools and none of the studies involved secondary school (Year 7–10) students. Moreover, as the meta-analysis was focused on mathematics education studies from countries where the language of instruction was English, the analysis of studies seems to be dominated by research from the USA with one study from Australia. Interestingly, no New Zealand-based studies were found in the meta-analysis. This lack of quantitative research does raise questions about the nature of research being conducted in New Zealand.

The findings suggested four main intervention categories, which are: Dual Language programme, Professional Development for Teachers, Curriculum Intervention, and Cognitively Focused Interventions.

Figure 6 provides a summary of interventions as successful teaching practices for teaching and learning of English Language Learners and their effect sizes.

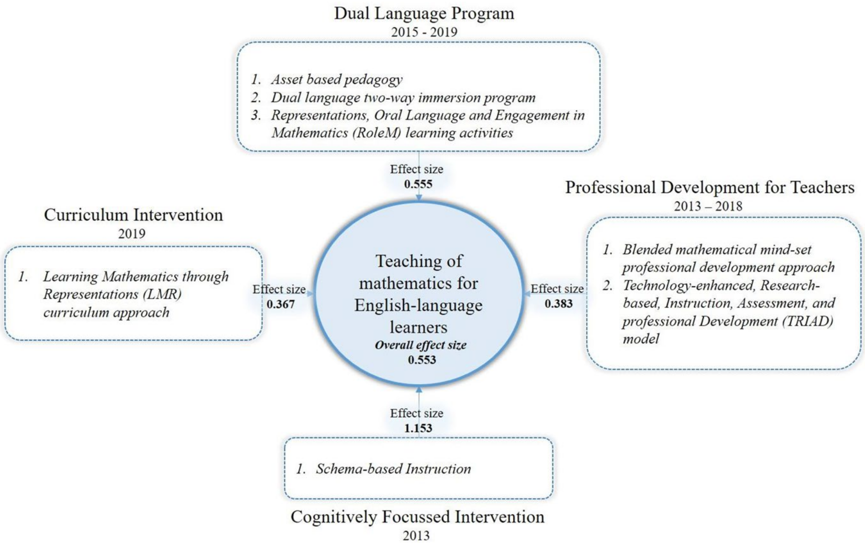


Fig. 6 Successful Teaching Practices and their Effect Sizes

Of the four different categories of interventions focused on supporting the mathematics learning of ELLs, the Cognitively Focused Interventions reported the largest effect size (1.153), followed by the Dual Language Programme interventions with an effect size of 0.555. It needs to be noted that the effect size of Dual Language Programme interventions is calculated based on a sample size of 3308 in comparison with sample size of 136 for Cognitively Focused Interventions. Hence, Dual Language programme interventions may hold better potential than Cognitively Focused Interventions (Marín-Martínez & Sánchez-Meca, 2010). Within the set of three Dual Language Programme studies the two-way immersion intervention reported by Vela et al. (2017) stood out as having a large individual effect size (0.975). Interventions involving professional development for teachers and the one study categorised as a Curriculum Intervention reported moderate effect sizes. Overall, the meta-analysis indicates the value of interventions that focus on supporting students' home languages alongside the development of the mathematical and English language.

Unsurprisingly, most of the successful teaching practices interventions related to mathematics education for ELLs involved Dual Language Programmes with a combined effect size of 0.555 with 95% CI. These teaching practices are: (i) asset-based pedagogy (Matthews & López, 2019), (ii) Dual language two-way immersion programme (Vela et al., 2017), and (iii) Representations, Oral Language and Engagement in Mathematics (RoleM) Learning activities (Warren & Miller, 2015). The effect size of dual language two-way intervention detailed by Vela et al. (2017) was 0.975, which is high. In this intervention, students were taught content and literacy in two languages. The findings concur with the conclusions of Lindholm-Leary (2013) who also claimed dual language two-way programmes are more effective than transitional bilingual education programmes, yet not all schools offer dual language two-way programmes for their students. The effect sizes from RoleM activities (Warren & Miller, 2015) were moderate with 0.688 for Year 1 and 0.507 for Year 2 students. Warren and Miller (2015) claimed that RoleM learning starts with exploring one concept using one representation, moving to using two representations in parallel, linking the parallel representations, and finally integrating representations. From their perspective, code switching relates to switching between representations, and code mixing relates to changing the oral language used to assist in making connections between the representations. Bose and Choudhury (2010) have also supported the use of both code switching and code mixing are important in breaking down the language barrier. Warren and Miller (2015) further contend that communicating mathematically involves aspects of both code switching and code mixing, they claim that these aspects are more than simply translating from one language to as suggested by Farrugia (2009). The third successful teaching strategy within the intervention category of Dual Language Programmes is the asset-based pedagogy (Matthews & López, 2019), which integrated aspects of cultural content and heritage language into classroom instruction. The individual effect size for this study is 0.140. The findings from Matthews and López (2019) indicate that integrating students' culture into instruction through honouring their home language is fundamental to student learning, even in mathematics where the connections between bilingualism and learning may be less clear. While research suggests that teacher expectations (Ladson-Billings, 1995; Turner et al., 2015) can directly and positively predicted mathematics achievement, Matthews and López (2019) study reveals that simply having high expectations for historically marginalised English Language Learners is not enough for teacher enactment of asset-based pedagogies. This may suggest that teachers with high expectations that are rooted in deficit, culturally biased thinking may subconsciously prioritise English instruction and thus subjugate students' heritage language. The findings resonate with Creese and Blackledge (2010) who write that the teachers need to go beyond acceptance or tolerance of children's languages to using language flexibly and encouraging the use of learners' linguistic repertoires. Tshabalala and Clarkson (2016) remind us that a teacher's use of learner's home language is not always effective. Confusion and misconceptions in teaching can arise if the teacher is not proficient in the home language of the learners or in the English mathematical language that was the focus of the teaching.

The effectiveness of Dual Language Programmes highlighted in this meta-analysis concurs with the earlier meta-analysis conducted by Reljić et al. (2015). They indicated that bilingual programmes have better outcomes than submersion programmes using results from a random-effect model of five studies. The findings of the present meta-analysis are also in agreement with those of previous meta-analyses (Krashen & McField, 2005; Rolstad et al., 2008; Slavin & Cheung, 2005) in favouring dual language or bilingual education that considers the use of the home language of children as a resource in the teaching and learning of school subjects to promote their academic achievement. Overall, these programmes have a simultaneous focus on language development and content development and argue that the instruction and student contributions can be in both languages.

As stated earlier, the effect size of Cognitively Focused Interventions (CFI) was the highest and the intervention is mentioned in one of the seven studies selected for this meta-analysis. The teaching practice is Schema-based Instruction (SBI) (Jitendra et al., 2013) with (ES 1.153). Jitendra et al. (2013) argued that SBI is effective for helping students to learn explicit word problem-solving procedures. They suggest that SBI support self-regulation when finding solutions and reduce working memory load by allowing students to break the problem-solving process into a sequence of manageable-sized tasks. The improvement in SBI student's problem-solving performance is similar to the medium to large effect sizes found in previous studies on grade 3 students (Fuchs et al., 2008; Jitendra et al., 1998). Cognitively Focused Interventions employs a structured sequence of steps with a teacher focused on ensuring that students (including ELLs) understand and can use the mathematical ideas that are the focus of teaching. The intervention approaches involve explicit instruction accompanied by a direct focus on metacognition—teachers prompt students to think about the processes they are using while students are working through the different steps of problem-solving. The labelling and nature of the steps in the different interventions was not the same but each intervention involved students moving through a sequence of understanding the problem and its components before developing a plan to solve the problem. Students then worked through their plan and evaluated their solutions. Teachers began the instructional sequences using a range of scaffolds which they gradually faded as students gained competence and confidence. The study made deliberate use of a range of visual, virtual and material representations of ideas. Students were guided in the use of these as resources for their thinking. The resources acted to reduce the cognitive load of a problem while providing students with access to ideas and resources they could use in the future.

In this meta-analysis, a teaching practice based on Curriculum Intervention was also identified, which was Learning Mathematics through Representations (LMR) curriculum (Saxe & Sussman, 2019) with effect size of 0.367. The LMR curriculum focused on a productive mathematical practice of representation using a design-based research approach. The intervention focused on integers and fractions; however, students in LMR showed marked gains on a general assessment of maths proficiency (the state standardised assessment). One reason for this is that LMR teachers developed more inclusive instructional practices and then sustained their use of LMR design principles through the remainder of the year. Another possibility is that the gains that ELLs made during the intervention phase seeded a developmental process that enhanced some ELLs' ability to engage with mathematics in other domains and with other materials. Saxe and Sussman (2019) argued that students developed higher-order thinking skills such as logical reasoning, mathematical thinking, and positive mathematics learning beliefs and behaviours as they learned through this curriculum approach. Similar to other design-based research projects (Cobb et al., 2003), the five-phase LMR lesson affords teachers opportunities to assess and integrate student reasoning in discussions and to adapt their instruction as students with diverse understandings and linguistic proficiencies reason publicly with varied representational formats. Saxe and Sussman (2019) argue that the use of representations and the lesson sequence increases the probability that all students, including ELLs who may have trouble accessing conventional mathematics curricula, are provided opportunities to engage with complex mathematical ideas such as fractions and integers. The intervention had moderate effect size, which may suggest that interventions with focus on overall curriculum and not just pedagogy have the potential to improve student outcomes. The findings, methods and design-based research approach will be of value for researchers and professionals developing teaching practices that engage all children with rich learning opportunities.

Only two of the seven studies provided two successful teaching practices with direct focus on professional development for teachers. These teaching practices are: (i) Blended mathematical mindset professional development approach (Anderson et al., 2018) and (ii) Technology-enhanced, Research-based, Instruction, Assessment, and professional Development (TRIAD) model (Clements et al., 2013). Anderson et al. (2018) focused on changing teachers' perceptions of their students and who could achieve in mathematics that is on promoting a shift in teacher practices from a fixed mindset to growth mindset approach. The online professional development course and the network meetings provided teachers multiple opportunities to struggle with and reflect on their deeply held beliefs about themselves and about mathematics. Anderson et al. (2018) reported an effect size of 0.437 for student achievement. The results of the Mathematical Mindset approach are noteworthy with the students most in need of support including girls, language learners, and economically disadvantaged students, changing their ideas and their achievement most significantly. At the heart of the student and the teacher, change was a change in mindset, in beliefs about learning, and the eradication of learning myths that have held back generations of mathematics learners. The data from this study add weight to a growing realisation that changing the myth that people are born as "math people" or not is one of the most important responsibilities for educators, now and in the future. Previous research on professional development (Borko et al., 2010) has recommended that professional development needs to be face-to-face, long term, and focused on content standards. This study shows the impact of an online class which is focused less upon standards and more upon personal growth, mindset, and belief in the potential of all learners. Similarly, Clements et al. (2013) worked with teachers to help them incorporate research-based ideas about student mathematical thinking and possible/productive learning trajectories into their practice. They reported an effect size of 0.355. Clements et al. (2013) suggest that the maintenance of the effect size in the TRIAD Follow-Through intervention, compared to the decreasing effect size in the Non-Follow-Through (TRIAD-NFT) intervention. Similar findings were reported by McLoyd (1998) who argues that centring teaching around learning trajectories may focus teachers' attention on students' thinking and learning of mathematics rather than their memberships in ethnic groups. As a result, this changed focus helped avoid perceptions that negatively affect teaching and learning.

Together these two studies suggest that successful teaching for English Language learners benefits from teachers reconsidering their view of these students and adjusting their teaching practices to take account of and support student learning potential, pathways, and understanding of what it means to do mathematics. Overall, this aspect of the studies reviewed here raises questions about the need for teacher professional learning to target the strengths and needs of particular groups/groupings of students in order to provide equal access to high quality mathematics learning opportunities.

Both these studies provide examples of what is possible when researchers work with practitioners and study multiple forms of data and change.

It is surprising that none of the studies identified in this review focused on enhancing statistical literacy of English Language Learners. In recognition of the importance of statistics in both school and out of school settings, there is a movement in many countries to include statistics at every level in the mathematics curricula (Lesser et al., 2016). Statistics and mathematics share commonalities, but require different thinking (Marshman et al., 2015), and use different languages (Dunn et al., 2016). For example, in mathematics, results are usually reached by means of deduction, logical proof, or mathematical induction and typically there is one correct answer. Statistics, however, utilises inductive reasoning and conclusions are always uncertain. Hence practices identified for mathematics teaching may not apply to statistics teaching. By drawing on research from across schools research can clarify these differences and give advice for mathematics teachers teaching statistics in schools. This could be an area of future meta-analysis and research.

Limitations of the meta-analysis

Meta-analysis aims to thoroughly examine the empirical evidence available on a certain topic, with this evidence included or not in the analysis based on a fixed set of criteria. However, this meta-analysis has a few limitations.

First, in this meta-analysis, quantitative studies on student achievement from which an effect size could be derived were included. Therefore, quantitative studies without relevant statistical data (such as partial correlation coefficients in case of multiple regression analysis) could not be included as their regression analysis would have examined the impact of more than one independent variable on a dependent variable (in this case, students' mathematics performance), for example, Han (2012).

Second, the studies published in English were selected for this meta-analysis, which may have resulted in the exclusion of potential studies on ELLs experiences published in other languages. The selection criteria also highlight the publication bias that needs to be taken into consideration while using the results mentioned in this paper.

Third, the focus on quantitative studies led to the exclusion of the considerable body of qualitative work by mathematics educators including that by Barwell et al. (2016), Barwell (2020), Moschkovich (2018), and Sharma (2016), among others, that acted as selection bias in this meta-analysis.

Fourth, studies based in Canada, USA, Australia, Scotland, Ireland, England, Wales, and New Zealand were selected for this meta-analysis. Hence, research in countries such as South Africa, China, Korea among others were not included. For example, Prediger and Wessel (2013) study from Germany was excluded in this meta-analysis.

Finally, given some of the included studies involved a reasonably small number of students. This is one of the main critiques of meta-analyses is that they try to compare incomparable elements (Borenstein et al., 2009). In this meta-analysis, relatively extended descriptions of the included studies are provided to mitigate this drawback.

Recommendations

The following recommendations for resourcing and support, and research are suggested to respond to the findings from this review.

Resourcing and support

- Assist teachers to explore the use of schema-based instruction based on principles of schema theory and complemented with think aloud protocols for enhancing students' performance on word problem-solving.
- Develop a version of the Dual Language Programmes identified in the study such as asset-based pedagogy (Matthews & López, 2019), a two-way immersion programme (Vela et al., 2017) and RoleM learning activities (Warren & Miller, 2015), with any development informed by Tātaiako and the Pacific Education Plan.
- Explicitly focus on the introduction and use of a range of representations building on, but not limited to, ideas described in RoleM learning activities (Warren & Miller, 2015) and LMR curriculum (Saxe & Sussman, 2019). Offer professional development that assists and provides time for teachers to reflect on their mathematical mindsets (Anderson et al., 2018) and reconsider their practices based on the evidence

Research

- Consider funding quantitative/mixed-method studies related to successful mathematical practices for English Language Learners in New Zealand and Australian context.
- Interestingly, all the studies identified in the meta-analysis focused on learning of ELLs in primary school settings. Research is required with focus on secondary mathematics education (Year 8–10).
- In the studies identified in this research, the focus seems to be on either overall mathematics achievement or numeracy concepts, and more quantitative research is required to develop an understanding of ELLs learning of statistics concepts.
- We suggest that research synthesising the qualitative research in this field would help in developing cumulative understanding of best teaching practices as suggested by Barwell (2020) and de Araujo et al. (2018).

Conclusion

In this study, a meta-analysis of quantitative and mixed-method studies from 2009 to 2019 was conducted to inform statistically informed successful teaching practices for improving learning outcomes for English Language Learners. Like proponents of the language-as-resource perspective, the meta-analysis suggests that students' home languages can be used as resources and may result in statistically significant teaching practices. It is to be noted that equity and academic excellence will not be attained until learners' home language is used as a resource in multilingual classrooms. In fact, bilingual resources and teaching practices directed towards ELLs in multicultural settings benefits all students who struggle with language use in mathematics. This view has implications for our Education policy (Ministry of Education, 2007) which states that all learners need to feel secure in their identities, languages and cultures and contribute fully to Aotearoa New Zealand's social, and cultural well-being.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This study was funded by the New Zealand Ministry of Education.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Studies included in the meta-analysis are indicated by the **

**Anderson, R. K., Boaler, J., & Dieckmann, J. A. (2018). Achieving elusive teacher change through challenging myths about learning: a blended approach. *Education Sciences, 8*(3), 98. https://doi.org/10.3390/educsci8030098

Anthony, G., & Walshaw, M. (2007). *Effective pedagogy in Pāngarau/Mathematics: Best evidence synthesis iteration (BES)*. Ministry of Education. https://www.educationcounts.govt.nz/publications/series/2515/5951

Barwell, R. (2020). Learning mathematics in a second language: Language positive and language neutral classrooms. *Journal for Research in Mathematics Education, 51*(2), 150–178. https://doi.org/10.5951/jresematheduc-2020-0018

Barwell, R. (Ed.). (2009). *Multilingualism in mathematics classrooms—global perspectives.* Multilingual Matters.

Barwell, R., Clarkson, P., Halai, A., Kazima, M., Moschkovich, J., Planas, N., Setati-Phakeng, M., Valero, P., & Villavicencio Ubillús, M. (2016). Introduction: an ICMI study on language diversity in mathematics education. In R. Barwell, P. Clarkson, A. Halai, M. Kazima, J. Moschkovich, N. Planas, M. Setati-Phakeng, P. Valero, & M. Villavicencio Ubillús (Eds.), *Mathematics education and language diversity: The 21st ICMI study* (1st ed. pp. 1–22). Springer International Publishing.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Bose, A., & Choudhury, M. (2010). Language negotiation in a multilingual mathematics classroom: an analysis. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education: Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia* (pp. 93–100). MERGA. https://eric.ed.gov/?id=ED520854

Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. In: P. Peterson, E. Baker, B. McGaw (Eds.), *International Encyclopedia of Education* (Vol. 7, pp. 548–556). Elsevier.

Byun, J., & Joung, E. (2018). Digital game-based learning for K–12 mathematics education: a meta-analysis. *School Science and Mathematics*, *118* (3–4), 113–126. https://onlinelibrary.wiley.com/doi/pdf/10.1111/ssm.12271

Clarkson, P. (2007). Australian Vietnamese students learning mathematics: High ability bilinguals and their use of their languages. *Educational Studies in Mathematics*, *64*(2), 191–215. https://link.springer.com/content/pdf/10.1007/s10649-006-4696-5.pdf

**Clements, D., Sarama, J., Wolfe, C., & Spitler, M. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies: Persistence of effects in the third year. *American Educational Research Journal, 50*(4), 812–850. https://doi.org/10.3102/0002831212469270

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13. https://journals.sagepub.com/doi/pdf/10.3102/0013189X032001009

Coldberg, P., & Bush, W. (2003). Using metacognitive skills to improve 3rd graders' math problem solving. *Focus on Learning Problems in Mathematics, 5*(10), 29–48.

Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach*. Sage.

Creese, A., & Blackledge, A. (2010). Translanguaging in the bilingual classroom: a pedagogy for learning and teaching? *The Modern Language Journal*, *94*(1), 103–115. https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-4781.2009.00986.x

Davenport, L. L., Kao, V. S., Matlen, B. J., & Schneider, S. A. (2019). Cognition research in practice: Engineering and evaluating a middle school math curriculum. *The Journal of Experimental Education*. https://doi.org/10.1080/00220973.2019.1619067

de Araujo, Z., Roberts, S. A., Willey, C., & Zahner, W. (2018). English learners in K–12 mathematics education: A review of the literature. *Review of Educational Research*, *88*(6), 879–919. https://journals.sagepub.com/doi/pdf/10.3102/0034654318798093

de Oliveira, L. C., & Cheng, D. (2011). Language and the multisemiotic nature of mathematics. *Reading Matrix: an International Online Journal, 11*(3), 255–268.

Dunn, P. K., Carey, M. D., Richardson, A. M., & McDonald, C. (2016). Learning the language of statistics: Challenges and teaching approaches. *Statistical Education Research Journal, 15*(1), 8–27.

Dweck, C. (2006). *Mindset: The new psychology of success.* Random House.

Education Review Office. (2018). Responding to language diversity in Auckland. The child the heart of the matter. https://ero.govt.nz/sites/default/files/2021-05/ALD-report2.pdf

Ellis, P. (2010). *The essential guide to effect sizes, statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.

European Commission. (2015). *Language teaching and learning in multilingual classrooms. Education and training brief.* Retrieved from http://ec.europa.eu/dgs/education_culture/repository/languages/library/policy/policy-brief_en.pdf

Farrugia, M. T. (2009). Reflections on a medium of instruction policy for mathematics in Malta. In R. Barwell (Ed.), *Multilingualism in Mathematics Classrooms: Global Perspectives* (pp. 97–112). Multilingual Matters.

Frigo, T., & Simpson, L. (2001). Research into the numeracy development of aboriginal students: Implications for the NSW K–10 mathematics syllabus. *Indigenous Education Research*, *8*. https://research.acer.edu.au/indigenous_education/8/

Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children*, *74*(2), 155–173. https://journals.sagepub.com/doi/pdf/10.1177/001440290807400202

García, O., & Leiva, C. (2014). Theorizing and enacting translanguaging for social justice. In A. Blackledge & A. Creese (Eds.), *Heteroglossia as practice and pedagogy* (pp. 199–216). Springer.

Hak, T., Van Rhee, H. J., & Suurmond, R. (2016). How to interpret results of meta-analysis (Version 1.3). The Netherlands. www.erim.eur.nl/researchsupport/meta-essentials/downloads

Halmos, P. R. (1980). The heart of mathematics. *The American Mathematical Monthly*, *87*(7), 519–524. https://www.jstor.org/stable/pdf/2321415.pdf

Han, W. (2012). Bilingualism and academic achievement. *Child Development*, *83*(1), 300–321. https://www.jstor.org/stable/41416083

Hipkins, R. (2019). Thinking critically about PISA. *Set: Research Information for Teachers*, *2*, 49–52. https://doi.org/10.18296/set.0143

Jackson, K., & Cobb, P. (2010). *Refining a vision of ambitious mathematics instruction to address issues of equity* [Paper presentation]. The Annual Meeting of the American Educational Research Association. Denver.

Jitendra, A. K., Griffin, C., McGoey, K., Gardill, C., Bhat, P., & Riley, T. (1998). Effects of mathematical word problem solving by students at risk or with mild disabilities. *Journal of Educational Research*, *91*, 345–356. https://www.tandfonline.com/doi/pdf/10.1080/00220679809597564

**Jitendra, A., Rodriguez, M., Kanive, R., Huang, J., Church, C., Corroy, K., & Zaslofsky, A. (2013). Impact of small-group tutoring interventions on the mathematical problem solving and achievement of third-grade students with mathematics difficulties. *Learning Disability Quarterly*, *36*(1), 21–35. https://www.jstor.org/stable/24570131

Kazima, M. (2006). Malawian students meaning for probability vocabulary. *Educational Studies in Mathematics, 64*(2), 169–189. https://doi.org/10.1007/s10649-006-9032-6

Krashen, S., & McField, G. (2005). What works? Reviewing the latest evidence on bilingual education. *Language Learner, 1*(2), 7–10.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 47*, 465–491. https://doi.org/10.3102/00028312032003465

Lesser, L. M., Wagler, A. E., & Salazar, B. (2016). Flipping between languages? An exploratory analysis of the usage by Spanish-speaking English language learner tertiary students of a bilingual probability applet. *Statistics Education Research Journal*, *15*(2), 145–168. https://doi.org/10.52041/serj.v15i2.245

Lesser, L. M., & Windsor, M. S. (2009). English language learners in introductory statistics: Lessons learned from an exploratory case study of two pre-service teachers. *Statistics Education Research Journal*, *8*(2), 5–32. https://bit.ly/2XEuUDg

Lindholm-Leary, K. (2013). Bilingual and biliteracy skills in young spanish-speaking low-SES children: Impact of instructional language and primary language proficiency. *International Journal of Bilingual Education and Bilingualism, 17*(2), 144–159. https://doi.org/10.1080/13670050.2013.866625

Lovakov, A., & Agadullina, E. (2021). Empirically derived guidelines for interpreting effect size in social psychology. *European Journal of Social Psychology., 51*(3), 458–504. https://doi.org/10.1002/ejsp.2752

Marín-Martínez, F., & Sánchez-Meca, J. (2010). Weighting by inverse variance or by sample size in random-effects meta-analysis. *Educational and Psychological Measurement*, *70*(1), 56–73. https://journals.sagepub.com/doi/pdf/10.1177/0013164409344534

Marshman, M., Dunn, P. K., McDougall, R., & Wiegand, A. (2015). A case study of the attitudes and preparedness of secondary mathematics teachers towards statistics. *Australian Senior Mathematics Journal*, *29*(1), 51–64. https://search.informit.org/doi/pdf/10.3316/informit.334335533701614

**Matthews, J. S., & López, F. (2019). Speaking their language: the role of cultural content integration and heritage language for academic achievement among Latino children. *Contemporary Educational Psychology, 57*, 72–86. https://doi.org/10.1016/j.cedpsych.2018.01.005

McLoyd, V. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53*, 185–204.

Ministry of Education. (2007). *The New Zealand curriculum.* Learning Media.

Ministry of Education. (2019). *PISA2018 New Zealand summary report system performance & equity.* New Zealand. https://www.educationcounts.govt.nz/__data/assets/pdf_file/0006/196629/PISA-2018-NZ-Summary-Report.pdf

Ministry of Education. (2014). *Ministry of Education Statement of Intent, 2014–2018*. New Zealand. https://www.education.govt.nz/assets/Documents/Ministry/Publications/Statements-of-intent/2014SOI.pdf

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine, 6*(7), 264–270. https://doi.org/10.1371/journal.pmed.1000097

Moschkovich, J. (2018). Talking to learn mathematics with understanding: Supporting academic literacy in mathematics for English learners. In A. Bailey, C. Maher, & L. C. Wilkinson (Eds.), *Language, literacy, and learning in the STEM disciplines: Language counts for English learners* (pp. 13–34). Routledge.

Moses, R. P., & Cobb, C. E. (2001). *Radical equations: Civil rights from Mississippi to the algebra project.* Beacon Press.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. http://timssandpirls.bc.edu/timss2015/international-results/

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.

Planas, N., & Setati-Phakeng, M. (2014). On the process of gaining language as resource in mathematics education. *ZDM, 46*, 883–893. https://doi.org/10.1007/s11858-014-0610-2

Prediger, S., & Wessel, L. (2013). Fostering German-language learners' constructions of meanings for fractions—design and effects of a language-and mathematics-integrated intervention. *Mathematics Education Research Journal, 25*(3), 435–456. https://doi.org/10.1007/s13394-013-0079-2

Reljić, G., Ferring, D., & Martin, R. (2015). A meta-analysis on the effectiveness of bilingual programs in Europe. *Review of Educational Research, 85*(1), 92–128. https://doi.org/10.3102/0034654314548514

RNZ. (2020). NZ students record worst results in maths and science. https://www.rnz.co.nz/news/national/432451/nz-students-record-worst-results-in-maths-and-science

Rolstad, K., Mahoney, K., & Glass, G. (2008). The big picture in bilingual education: a meta-analysis corrected for Gersten's coding error. *Journal of Educational Research & Policy Studies*, *8*(2), 1–15. https://files.eric.ed.gov/fulltext/EJ859089.pdf

Sarama, J., Clements, D. H., & Henry, J. J. (1998). Network of influences in an implementation of a mathematics curriculum innovation. *International Journal of Computers for Mathematical Learning, 3*, 113–148. https://doi.org/10.1023/A:1009768005398

Sarama, J., Clements, D. H., Wolfe, C. B., & Spitler, M. E. (2012). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies. *Journal of Research on Educational Effectiveness, 5*(2), 105–135. https://doi.org/10.1080/19345747.2011.627980

**Saxe, G. B., & Sussman, J. (2019). Mathematics learning in language inclusive classrooms: Supporting the achievement of English learners and their English proficient peers. *Educational Researcher, 48*(7), 452–465. https://doi.org/10.3102/0013189X19869953

Schleppegrell, M. J. (2011). Language in mathematics teaching and learning: a research review. In M. J. Schleppergrell (Ed.), *Language and mathematics education: Multiple perspectives and directions for research* (pp. 197–216). Continuum.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist, 47*(10), 1173–1181.

Sharma, S. (2016). Language barriers in statistics education: Some findings from Fiji. *International Journal of Learning, Teaching, and Educational Research, 15*(8), 23–34.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75*, 247–284. https://doi.org/10.3102/00346543075002247

The Partnership for 21st Century Skills. (2009). The intellectual and policy foundations of the 21st century skills framework. http://route21.p21.org/images/stories/epapers/skills_foundations_final.pdf

Tran, N. A., Behseta, S., Ellis, M., Martinez-Cruz, A., & Contreras, J. (2015). The effects of spanish english dual language immersion on student achievement in science and mathematics. *EJEP: EJournal of Education Policy*, 57–77.

Tshabalala, L., & Clarkson, P. (2016). Complementary functions of learning mathematics in complementary schools. In A. Halai & P. Clarkson (Eds.), *Teaching and learning mathematics in multilingual classrooms* (pp. 227–247). Brill Sense Publishers. https://brill.com/view/book/edcoll/9789463002295/BP000016.xml

Turner, H., Rubie-Davies, C. M., & Webber, M. (2015). Teacher expectations, ethnicity and the achievement gap. *New Zealand Journal of Educational Studies, 50*(1), 55–69. https://doi.org/10.1007/s40841-015-0004-1

**Vela, A., Jones, D., Mundy, M. A., & Isaacson, C. (2017). Determining the effectiveness of bilingual programs on third grade state exam scores. *Research in Higher Education Journal*, *33*. https://www.aabri.com/manuscripts/172633.pdf

Warren, E., & deVries, E. (2009). Young Australian Indigenous students' engagement with numeracy: Actions that assist to bridge the gap. *Australian Journal of Education, 53*(2), 159–175. https://doi.org/10.1177/000494410905300205

**Warren, E., & Miller, J. (2015). Supporting English second-language learners in disadvantaged contexts: Learning approaches that promote success in mathematics. *International Journal of Early Years Education, 23*(2), 192–208. https://doi.org/10.1080/09669760.2014.969200