



Self-efficacy in teaching mathematics and the use of effective pedagogical practices in New Zealand primary schools

David A. G. Berg¹ · Naomi Ingram¹ · Mustafa Asil² · Jenny Ward¹ · Jeffrey K. Smith¹

Accepted: 23 January 2024
© The Author(s) 2024

Abstract

This study explored teachers' self-efficacy in teaching mathematics (SETM) as related to their teaching profile and pedagogical practices. Using data from 327 New Zealand primary teachers, a multilevel structural equation model was constructed and analyzed that looked at the relationships among SETM and effective pedagogical practice scales and included characteristics related to the teachers and their schools. No significant gender, year level, or school socioeconomic differences were found with teachers' SETM. Experienced teachers had significantly higher SETM scores than their inexperienced colleagues. Teachers who were self-efficacious in their mathematics teaching reported pedagogical practices known to be effective in the mathematics classroom in contrast to their low self-efficacy colleagues.

Keywords Teacher self-efficacy · Self-efficacy in teaching mathematics (SETM) · Primary teaching · Affect · Effective pedagogy · Multilevel structural equation modeling (SEM)

✉ David A. G. Berg
david.berg@otago.ac.nz

Naomi Ingram
naomi.ingram@otago.ac.nz

Mustafa Asil
masil@bond.edu.au

Jenny Ward
jenny.ward@otago.ac.nz

Jeffrey K. Smith
jeffrey.smith@otago.ac.nz

¹ College of Education, University of Otago, Dunedin, New Zealand

² Faculty of Health Sciences and Medicine, Bond University, Robina, Australia

Introduction

Success in mathematics has been found to be dependent on cognitive factors such as intelligence, spatial ability, problem solving skills, and mastery of content (Hilbert et al., 2019; Khine, 2017), as well as the pedagogical practices used by teachers of mathematics, such as building on students' thinking, and providing challenging mathematics problems (Anthony & Walshaw, 2009a). The affective aspects of learning and teaching have also been recognized as important for success in mathematics and related subjects (Beatson et al., 2018; Grootenboer & Marshman, 2016; Ingram et al., 2020; Lee & Shute, 2010; Eynde et al., 2006). Included within these studies of affect is a significant body of work on teachers' beliefs about mathematics, and their pedagogical practices (Attard et al., 2016; Murphy & Ingram, 2023), as well as teacher self-efficacy (TSE) (e.g., Carney et al., 2016), and self-efficacy in teaching mathematics (SETM) (Bjerke & Xenofontos, 2023).

The purpose of this research is to explore the pedagogical practices, as well as the teacher and school characteristics, of teachers who report different levels of SETM in a nationally representative sample. More specifically, we consider the relationship between SETM and the self-report of teaching practices associated with teaching effectively in mathematics. Additionally, we explore differences in SETM by gender, level taught, experience, and school socioeconomic context.

Central to our approach to exploring teachers' SETM is the use of Bandura's (1997) social cognitive theory (SCT). Bandura (1997) argued that "perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 3).

Self-efficacy

In his development of social cognitive theory, Bandura (1986) challenged the behaviorist understanding of human action as being largely a reaction to environmental stimuli by recognizing that individuals can influence, and be influenced by, their environments and behaviors. In what he termed, "triadic reciprocal causation," behavior, personal factors, and environment are all related, as shown in Fig. 1.

A good example of this in an educational setting can be found in a study by Brouwers et al. (2001), where teachers experiencing burnout exhibited behaviors that shaped and were shaped by their classroom environments.

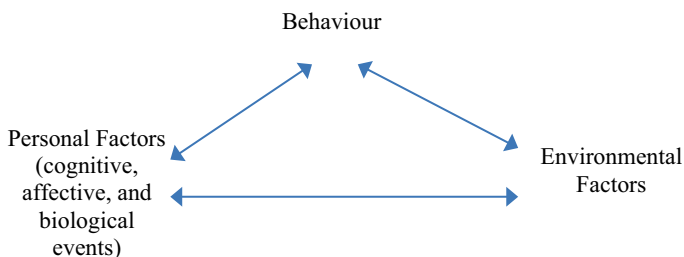


Fig. 1 Reciprocal causation: the triad of determinants in Bandura's social cognitive theory (Adapted from Bandura 1997, p. 6)

Within Bandura's social cognitive theory, self-efficacy is an individual's belief about the extent to which they have the necessary competency to overcome challenges to achieve a goal (Bandura, 1997). These beliefs answer questions such as "Can I do it?" and "How well can I do it?" (Skaalvik & Skaalvik, 2017, p. 153). Skaalvik and Skaalvik (2017) argue that self-efficacy should be distinguished from the related idea of academic self-concept, which asks, "Am I good at it?". Self-efficacy beliefs have been shown to exert enormous influence on people's lives. They govern the choices that they make about alternative courses of action; their effort, persistence, and resilience in pursuing results; the nature of their thought patterns; and the success of their endeavors (Bandura, 1997, 2006). Such choices are pertinent to career selection and success in careers (Bandura, 1997; Blotnicky et al., 2018), including motives to become a teacher (Berg et al., 2023).

Teacher self-efficacy

Drawing from Bandura's (1997) work, Skaalvik and Skaalvik (2010), defined Teacher Self Efficacy (TSE) as "...individual teachers' beliefs in their own ability to plan, organize, and carry out activities that are required to attain given educational goals" (p. 1059). In a body of work spanning four decades, strong TSE beliefs have been found to be correlated with a wide range of desired outcomes for teachers and their students, including student motivation (Zee & Koomen, 2016), academic achievement and interaction quality in mathematics (Perera & John, 2020), and higher levels of teacher job satisfaction and teacher engagement (Brouwers & Tomic, 2000; Collie et al., 2012; Klassen & Chiu, 2010; Klassen et al., 2013; Skaalvik & Skaalvik, 2007, 2010, 2017; Wolters & Daugherty, 2007).

Teacher self-efficacy and pedagogical practices

Of relevance to this present study, TSE beliefs have been found to be related to teachers' pedagogical practices (Holzberger et al., 2013; Tschannen-Moran et al., 1998). TSE researchers Guo et al. (2012) and Woolfolk Hoy and Davis (2005) have suggested that the relationship between TSE and student outcomes is indirect, with instruction and classroom environment mediating the relationship. Indeed, teachers with robust TSE beliefs have been found to use more innovative methods and new strategies (Chatzistamatiou et al., 2014; Cousins & Walker, 2000) and to set higher learning goals for learners (Wolters & Daugherty, 2007). Furthermore, TSE beliefs have also been found to (indirectly) predict positive engagement in professional learning (Durksen et al., 2017).

However, as with many studies of this type, determining the direction of causality is problematic. Zee and Koomen (2016) reviewed 25 research studies that claimed that the instructional behaviors and strategies teachers used to encourage learning may, in part, be influenced by their TSE. However, they noted that "around 70% of the 25 reviewed articles on instructional support relied on simple correlations and global measures of TSE, making it difficult to determine whether particular domains of TSE have similar patterns of effects on teachers' instructional support" (p. 990). These authors also raised concerns that over half of the studies they reviewed included samples of fewer than 100. Furthermore, Holzberger et al. (2013) examined the relationship between the TSE and instructional quality of 155 German secondary school mathematics teachers and found only a limited confirmation of the casual effect of TSE on instructional quality, but they did report an effect of instructional quality on TSE. Their research focused on student perceptions of the instruction quality actually delivered (as opposed to the instructional decisions that teachers take).

Their findings are consistent with Bandura's argument that mastery experiences influence self-efficacy beliefs (Bandura, 1997). The research we are looking at posits that higher TSE leads teachers to choose to use more effective instructional approaches. Over time, one would expect the use of more effective instructional approaches would lead to more effective instruction, and therefore consequently higher TSE. Or, as Bandura's (1993) theory suggests, the relationships are likely to ultimately be reciprocal. But in our work here, we are focused on TSE and the choices teachers take about how to deliver instruction.

Teacher/school characteristics and teacher self-efficacy

In our review of the literature, we were only able to locate a limited number of studies that explored the relationship between teacher/school characteristics and TSE belief. This is consistent with the experiences of other researchers, such as Klassen and Chiu (2010), who noted, "There has been surprisingly little research on how teaching context influences teachers' self-efficacy." For example, we were surprised to find a paucity of literature exploring the relationship between the socioeconomic status of school communities and TSE beliefs. Aligned with our experience, Malmberg et al. (2014) noted that there was "... a dearth of studies which have incorporated student group level or school-level variables of disadvantage as predictors of TSE" (p. 432).

Studies have contributed to our understanding of how teaching different grades of students may relate to TSE beliefs. Klassen and Chiu (2010) reported that compared to teaching in a secondary context, teaching in kindergarten and elementary schools was associated with higher self-efficacy related to student engagement and classroom management. They found this pattern repeated within teaching levels, where teachers of the lowest grades reported the highest self-efficacy for student engagement and classroom management. Interestingly, this was not the case for the use of instructional strategies. These findings are well aligned to other studies that also reported an inverse relationship between grade level and TSE (e.g., Durksen et al. 2017; Wolters and Daugherty 2007).

Studies that considered the relationship between teachers' years of service and their TSE beliefs have yielded inconsistent results (Ghaith & Yaghi, 1997). Some studies found that more seasoned teachers reported higher self-efficacy for student support and organizing the classroom (Gabriele & Joram, 2007; Wolters & Daugherty, 2007), whereas others, such as Guo et al. (2010), in their study of early childhood teachers, found the opposite to be true. Skaalvik and Skaalvik's (2007) research on Norwegian elementary and middle school teachers found years of service to be negatively correlated with three of the six subscales from the Norwegian TSE Scale: motivating students, coping with changes and challenges, and cooperating with colleagues and parents. These apparent contradictions may be understood, in part, in the light of Klaasen and Chiu's (2010) assertion that TSE has been found to increase across the first two decades of teaching and then decrease toward the end of a teacher's career. They noted a pattern of growth up to 23 years and then a gradual decline in the areas of using effective instructional strategies, engaging students, and managing student behavior. Work still needs to be done to gain a more complete understanding of how and why these constructs change over time.

We found scant research on teacher gender and TSE, and those studies that did exist reported dissimilar findings from different contexts. For example, Raudenbush et al. (1992) reported that American male high school teachers had significantly lower level of TSE than female teachers, whereas Klassen and Chiu (2010) reported that the female K-12 teachers

in their Canadian study had lower self-efficacy in classroom management compared to male teachers.

Several confounding variables are likely to impact on studies of TSE, making generalization problematic. Such variables include the level of qualification and expertise in a subject, and the quality and availability of professional learning opportunities. However, even the contribution of professional learning opportunities to TSE may not be as straightforward as it may seem. Tschannen-Moran and McMaster (2009) found that professional learning experiences have the potential to elevate TSE, particularly those designed to provide authentic mastery experience. However, they also noted that those teachers with elevated TSE were more positive about their professional learning. Thus, the direction of influence is unclear, and likely to be reciprocal (Durksen et al., 2017).

Measurement issues for teacher self-efficacy

Overall, the body of TSE research has made a valuable contribution to the field of educational research, yet, as noted by prominent researchers in the field, measurement issues have been an ongoing concern. “A problem with research on TSE is that there is no common agreement about how the construct should be measured. It has been conceptualized and measured differently by different researchers” (Skaalvik & Skaalvik, 2010, p. 1). Zee and Koomen (2016) examined 25 survey studies exploring the relationship between TSE and instructional behavior. They raised important concerns about the “methodological choices” (p. 10) employed in many of them, stating that around 70% included small samples and “global measures” of TSE. Roberts and Henson (2001) raised concerns about “the construct validity of scores from a variety of instruments purporting to measure teacher efficacy ...” (p. 5). They argued that much TSE research was “theoretically confused and generally not reflective of Bandura’s (1986) social cognitive theory conceptualization of self-efficacy” (p. 6). These methodological difficulties are not unusual in the affective literature (Ingram, 2017); we return to this topic in relation to the definition of items used in our own study. One such methodological difficulty concerns the ubiquity of generalized measures of TSE. Teachers’ self-efficacy is not only context bound (Dellinger et al., 2008), but is also likely to be task bound and subject-specific (Philippou & Christou, 2002). Bandura (1997) argued that TSE measures should reflect this and be situated in specific learning areas. (This study offers a re-analysis of work conducted to look at student achievement in New Zealand. The instruments used in this study were taken from this more general study. It was not possible to administer a more extensive set of questionnaires expressly measuring these constructs; however, as discussed later, the instrument utilizes a large representative sample, is situated in a specific learning area: mathematics, and is bound within the context of New Zealand primary schooling.)

For comprehensive reviews and discussion of TSE literature, see Berg (2022); Klassen et al. (2011); Tschannen-Moran et al. (1998); and Zee and Koomen (2016).

Self-efficacy for teaching mathematics

There is an emerging body of research which explores TSE specifically related to the teaching of mathematics: self-efficacy in teaching mathematics (SETM). We contend that SETM is a domain-specific form of teacher self-efficacy (TSE), which is, in turn, a domain-specific form of self-efficacy. We use the term SETM and acknowledge this includes teachers’ self-efficacy in relation to their pedagogical practices when teaching mathematics and their

mathematics self-efficacy—their self-efficacy in relation to the discipline of mathematics (Bjerke & Xenofontos, 2023).

Our review of the literature suggests that there are a somewhat limited number of studies that focus on teachers' SETM, although we note the overall study of teachers' beliefs related to their mathematics teaching has decreased across the field of research in mathematics education (Ingram et al., 2020) and acknowledge a focus on preservice teachers' SETM (Bjerke & Xenofontos, 2023).

Despite the scarcity of research, it appears that SETM plays an important role in empowering teachers to respond where students' engagement in mathematics is low (Skillington et al., 2016). Furthermore, there is evidence that teachers' SETM is lower than TSE for engaging students in other areas of the curriculum, with many secondary mathematics teachers not investing in long-term efforts to motivate their students as they believe they lack the knowledge and skills to do so (Hardré, 2011). It is unclear whether this is because mathematics teachers lack strategies for engagement, or because mathematics engagement is perceived as being particularly challenging (Hardré, 2011). Lomas and Clarke (2016), despite noting the difficulty of changing an individual's beliefs about him or herself, found that teachers efficacy beliefs improved when they saw evidence that new approaches positively impact their students' mathematics learning (Lomas & Clarke, 2016).

Alongside this literature focused on SETM and mathematics engagement, there is very little work investigating the link between teachers' SETM for mathematics and their pedagogical practices, although we note a body of research that examines students' self-efficacy in relation to teachers' practices (Murphy & Ingram, 2023). For example, in a comprehensive review of the literature on the consequences of teachers' self-efficacy for the quality of classroom processes, Zee and Koomen (2016) found just two studies focused on teachers' behaviors and practices related to the provision of instructional support in mathematics. This lack of work is of note, because (as outlined above) teachers' self-efficacy is likely to be subject specific (Philippou & Christou, 2002) and effective measures of teacher self-efficacy need to link to particular knowledge domains (Bandura, 1997).

The small number of studies investigating how teachers' SETM affects their pedagogical practice suggest that teachers' efficacy beliefs for teaching mathematics are positively associated with interaction quality, a critical dimension of the quality of classroom processes (Perera & John, 2020). Furthermore, teachers' self-efficacy for carrying out and teaching problem solving has been found to affect teachers' reports of their use of student-centered practices and has also been found to increase in response to professional development programs (Saadati et al., 2021). While these findings are encouraging, self-efficacy is not always positively linked to student achievement as noted by Sarac and Aslan-Tutak (2017), who identified the need to better understand the relationship between teacher self-efficacy, classroom processes, and student achievement (Zee & Koomen, 2016).

Effective pedagogical practices in mathematics

There is debate about what constitutes effective pedagogical practice in mathematics. Often, models constructed to evaluate effective pedagogical practices have overlapping descriptions and lack common terminologies (Jacobs & Spangler, 2017). We contend that teachers' pedagogical practices can be evaluated according to how consistent they are with the kinds of teaching approaches known to have a positive impact on student learning (Anthony & Walshaw, 2009b). Generally, these teaching approaches include the creation of a supportive learning environment, where shared learning and making connections to

prior learning and relevant contexts are facilitated, where reflective thought is encouraged, where active learning takes place, and where teachers inquire into the impact of their own teaching (New Zealand Ministry of Education, 2007).

In this research, we used Anthony and Walshaw's (2009a) best evidence synthesis to underpin our understanding of effective pedagogy. These authors identified ten pedagogical practices that best support mathematical competencies and identities to develop. They argue that effective mathematics teachers have classroom communities where there is an ethic of care; students have opportunities to work independently and cooperatively, and actively participate in whole-class discussion; students' thinking is built upon; worthwhile mathematical tasks are provided; connections are made among mathematical concepts; a range of assessment practices are used; there are opportunities to communicate mathematically; mathematical language is used; tools and representations are carefully selected; and teachers have a sound grasp of the relevant content and pedagogical content knowledge. These principles are informed by evidence that these pedagogies lead to desirable academic outcomes for students, including conceptual understanding, procedural fluency, strategic competence, and adaptive reasoning (Anthony & Walshaw, 2009a).

This framework for effective pedagogy was written for an international audience but is relevant for the New Zealand context because of its focus on high expectations of mathematical teaching and learning, and its emphasis on respecting cultures that students bring to the classroom (Ingram, 2020), and its link to the pedagogical practices of the New Zealand curriculum (Ministry of Education, 2007). Despite being 15 years old, it remains a foundation document for mathematics education to inform policy intended to address concerns about New Zealand's students' mathematics achievement in national and international assessments (Ingram 2020).

Mathematics teaching and learning in New Zealand

Although this research is situated in the unique context of New Zealand, it is useful for international understanding of primary teachers' mathematics teacher self-efficacy. Like other countries (e.g., the USA, Australia, the United Kingdom), primary teachers in New Zealand rarely have specialist qualifications in mathematics and, in general, lack confidence in teaching mathematics, particularly to students aged 11–12 years (Cayhill & Zhao, 2020). Furthermore, similar to other countries, New Zealand has responded to comparatively poor results in international assessments with a variety of measures such as a focus on students' problem solving, the provision of challenging tasks, and a resulting focus on primary teachers' mathematics competency and mathematics teacher self-efficacy (Ingram, 2020; Russo et al, 2020).

Research questions

Measuring the SETM of teachers and understanding how SETM relates to their pedagogical practices and teacher/school characteristics can provide teacher educators, policy makers, principals, and teachers with important information related to improving teachers' teaching and students' learning. As described, studies have looked at various components of these issues, but New Zealand's National Monitoring Study of Student Achievement (Ministry of Education, 2015) provides a unique opportunity to look at these questions comprehensively as it uses a much larger sample size that is typically evident in the

literature. Furthermore, the sample is nationally representative and includes teachers from 100 schools at two primary levels. In this paper, our overarching research question is “How strongly associated are teacher/school characteristics and SETM with optimal teaching behaviors in primary level mathematics?” We develop and test a multilevel structural equation model that addresses this issue by testing the following hypotheses:

- *Hypothesis 1* Teachers’ SETM differs according to their teacher/school characteristics. (That is, what are the differences in SETM by year level taught, gender, school socioeconomic level, teaching experience, level of mathematics qualifications, and experience of professional learning and development?)
- *Hypothesis 2* Teachers’ teaching behaviors differ in primary mathematics according to their teacher/school characteristics. (That is, what are the differences in teaching quality by year level taught, gender, school socioeconomic level, teaching experience, level of mathematics qualifications, and experience of professional learning and development?)
- *Hypothesis 3* SETM is related to teaching behaviors in primary mathematics.

We note that our approach posits that SETM influences TB and not the other way around. As our data are collected at a single point in time, we cannot confirm nor disconfirm this hypothesis. Our approach here is based upon a logical argument of the situation. Specifically, it is our position that one’s SETM can influence the teaching choices that a teacher makes, such as choosing to use worthwhile mathematical tasks in one’s teaching, but it is basically illogical to think that choosing to teach in a specific fashion would then cause one to have higher SETM. Now, if making such a choice leads to a more successful year in teaching, then that positive mastery experience will likely lead to higher SETM. But just the act of taking a choice to change one’s teaching approach would not in and of itself lead to higher SETM. For example, if a teacher decides at the end of the school year to use worthwhile tasks, we would not expect their SETM to increase over the summer months. Such an increase would accompany a successful mastery experience in the classroom the following year. However, we acknowledge that this is a supposition on our part, and others may take a different view.

Method

Our study uses data collected from 327 teachers at 181 schools in New Zealand as part of the National Monitoring Study of Student Achievement: Wānangatia Te Putanga Tauira (NMSSA) Mathematics and Statistics Assessment. The wider NMSSA study collects achievement data, together with contextual and background information from Year 4 and Year 8 students (aged 8/9 and 12/13 years old respectively), teachers, and principals. Specially trained teacher assessors visit schools and use one-to-one interviews, hands-on activities; group tasks; pencil and paper tasks, and computer adapted activities to evaluate and understand student achievement across the learning areas of the New Zealand Curriculum at the primary (elementary) school level, without high stakes testing (Ministry of Education, 2015). One of the aims of NMSSA is to better understand the factors that influence student achievement at the middle and endpoints of primary schooling in New Zealand (Allan, 2012; Asil, 2017; Darr et al., 2017). It is important to note that primary schooling in New Zealand concludes at Year 8.

A representative sample of the teachers of the assessed students were asked to complete a questionnaire. The questionnaire included sections about their attitudes toward and confidence in teaching mathematics and statistics, the professional support they received for teaching this learning area, and the types of mathematics learning activities and experiences that they provided for their students. The questionnaire comprises a combination of selected-response, open-ended questions, and four-point Likert scales.

Participants

NMSSA employed a two-stage stratified sampling design to select nationally representative samples of students, and their teachers and principals. The teacher participants were 167 Year 4 and 160 Year 8 teachers from 181 New Zealand primary schools. Of the teachers, 250 (76.5%) were female and 67 (21.1%) were male. Ten teachers (3.1%) did not provide their gender. Representation by teaching experience was 10.0% (0–2 years), 8.1% (3–5 years), 25.0% (6–10 years), 33.8% (11–20 years), and 23.1% (more than 20 years). Of the teachers, 65 (19.9%) were working in low socioeconomic schools, 142 (43.4%) were working in mid socioeconomic schools, and 120 (36.7%) were working in high socioeconomic schools. The socioeconomic status of the schools was taken from the New Zealand Ministry of Education records. All participating teachers were trained as general classroom teachers. Fifteen of the Year 8 teachers (8.1%) and 7 of the Year 4 teachers (4.0%) reported working in specialist mathematics teacher roles. In the New Zealand primary teaching context, these specialist roles typically involve both working with small groups of students to accelerate progress and supporting classroom teachers to improve practice. These are leadership roles in addition to their generalist teaching responsibilities.

Measures

SETM scale

An instrument was constructed using NMSSA teacher questionnaire items to measure SETM. Items were selected from the teacher questionnaire that asked teachers to make a judgment about their capability to “execute certain types of performance” (Bandura, 2006, p. 309). Bandura (2006) has argued “The construction of sound efficacy scales relies on a good conceptual analysis of the relevant domain of functioning. Knowledge of the activity domain specifies which aspects of personal efficacy should be measured.” The NMSSA team comprises curriculum experts and primary teachers with extensive subject knowledge in mathematics in addition to psychometricians and educational psychologists. Thus, the NMSSA items used in our scale are underpinned by a strong conceptual understanding of the domain and its measurement. These items are all measured on a 4-point scale (1 = not at all true for me and 4 = very true for me). The SETM scale items we selected are provided in Table 1. The set of items was analyzed for internal consistency using confirmatory factor analysis (see Results). It is important to note that these items were not originally developed to form a SETM scale, but rather to provide useful information to the Ministry of Education. Thus, the wording of the items is not in complete alignment with recommendations made by Bandura (2006) for the construction of self-efficacy scales. Nevertheless, we argue that these items are indeed sufficient to act as a strong proxy for SETM, and given Zee and Koomen’s (2016) concerns about the small samples and “global measures” evident in studies of teachers’ self-efficacy, we believe the results and discussion from this

Table 1 Self-efficacy for teaching mathematics scale items

	Item statement
Item 1	I can gauge how well students understand what I have taught in maths
Item 2	I am confident about responding to difficult questions from my students in maths
Item 3	I can provide alternative explanations or examples for the students
Item 4	I am able to motivate students who show little interest in maths
Item 5	I feel confident about teaching maths
Item 6	I have the necessary knowledge and skills to teach maths to a diverse range of students

study using nationally representative data can make a valuable contribution to TSE/SETM research.

Effective pedagogical practice scale

The NMSSA teacher questionnaire also asked about teaching practices that participants employed in their teaching. Teachers were asked to report on the frequency they used specific practices within the mathematics classroom to develop students' learning. We grouped these items using the descriptions of effective pedagogical practices outlined in Anthony and Walshaw (2009a). Five of Anthony and Walshaw's (2009a) effective pedagogical practices were consistent with the practice items within the questionnaire and from four to six questionnaire items were identified for each effective pedagogy (see Table 2). For example, teachers who establish an ethic of care within their classroom develop caring classroom environments that have a strong mathematical focus and build students' mathematical identities. Students feel safe to critique ideas voiced by their classmates. So, items related to the learning environment, students' mathematical identities or understanding others' perspectives related to mathematical learning were grouped within an ethic of care.

Data analysis

Data were analyzed by Mplus 7 (Muthén & Muthén, 2012) and IBM-SPSS27. There were two stages of data analysis. Before attempting to answer the research questions, we needed to determine if there was evidence to support the factorial structure of the SETM scale and the five pedagogical practices scales.

When examining teachers within schools as a sample, it is anticipated that there will be a considerable amount of shared variance within each cluster. Failure to consider this factor may introduce bias to the standard error, potentially elevating the likelihood of a type I error. In the second stage, intraclass correlation coefficients (ICC) were estimated from the unconditional means models (null model with no predictors) to determine whether the nesting was to be included in the analyses (Raudenbush & Bryk, 2002). While there are no universally agreed-upon criteria for evaluating the value of intraclass ICCs, it has been proposed that ICCs equal to or greater 0.05 necessitate the implementation of multilevel modeling (MLM) (Geldhof et al., 2014).

After achieving satisfactory model-data fit for the measurement models, and examining variance in outcome measures due to clustering, a multilevel structural equation modeling

Table 2 Effective pedagogies for mathematics scales (adapted from Anthony and Walshaw 2009a)

Effective pedagogy	Description	Examples of items
An ethic of care (EC)	Effective teachers have caring and safe classroom communities where students are able to think, reason, communicate, reflect upon, and critique the mathematics they encounter. Students have positive mathematical identities	Take into consideration how other people might be feeling or thinking Talk about how they are learning and different ways of learning
Building on students' thinking (BT)	Teachers make decisions based on students' current knowledge and interests. Students can learn from their errors by the teacher organizing discussion where students share their interpretations or solution strategies so that they can compare and re-evaluate their thinking	Incorporate students' existing ideas and experiences into their learning Explore how new learning links to, or changes, what they already know
Worthwhile mathematical tasks (WT)	Tasks allow for original thinking about important concepts. Open-ended tasks foster the creative thinking and experimentation that characterize mathematical play	Work on challenging or interesting maths problems or ideas Work on maths problems or ideas suggested by students
Making connections (MK)	Teachers support students in creating connections between different ways of solving problems, between mathematical representations and topics, and between mathematics and everyday experiences	Talk about how students could use their learning in new situations Draw on students' background and experiences to support their learning in maths
Mathematical communication (MC)	Teachers encourage their students to explain and justify their solutions, to communicate their ideas orally, in writing, and by using a variety of representations. Students become accustomed to listening to the ideas of others and use debate to resolve conflict and arrive at common understandings	Critically review each other's work to give feedback Share solutions to a maths problem with other students

(SEM) approach was employed to investigate the relationships among SETM, effective teaching pedagogies (teaching behavior) and school and teacher characteristics.

Testing measurement models: evidence on the internal structure of the scales

We employed confirmatory factor analysis (CFA) to examine the internal factor structure of the SETM scale and pedagogical practices scales to provide validity evidence for their use. Weighted least squares mean and variance adjusted (WLSMV), which is a robust estimator designed for ordinal indicators (Flora & Curran, 2004; Sass et al., 2014), was applied. As suggested in the literature (Cheung & Rensvold, 2002; Fan & Sivo, 2005, 2007), we used a number of different indices to evaluate model fit. Following recommendations from the field (Brown, 2006; Browne & Cudeck, 1992; Hair et al., 2010; Hu & Bentler, 1998; MacCallum et al., 1996; Yu, 2002), we assessed goodness-of-fit using chi-square (χ^2), the comparative fit index (CFI), and the root-mean-square error of approximation (RMSEA) as indicators of “good” or “acceptable” fit. The cut-offs we adopted were a non-significant chi-square (or $\chi^2/df < 3$ due to the sensitivity of sample size), RMSEA with values < 0.08 being indicative of reasonable fit, and values < 0.05 indicating a good fit, CFI with values > 0.90 indicating an acceptable fit and values > 0.95 indicating a good fit.

Multilevel structural equation modeling (SEM): relationship among SETM, TB, and teacher/school characteristics

We began the analyses by first looking at how the teachers responded to the SETM questions in general. That is, we wanted to see how strong the teachers’ SETM was in an absolute fashion. Were they above or below the middle responses on the SETM scale? Then, we employed a multilevel SEM model (Fig. 2) with robust maximum likelihood (MLR) estimator to examine the complex relationships among the variables of interest.

In this model, we examined the relationships among SETM, teaching behavior (TB), and teacher/school characteristics. Those characteristics were year level taught (year 4 = 0, year 8 = 1), gender (male = 0, female = 1), school socioeconomic level (low, mid, high decile), teaching experience (less (0–5 years), mid (6–10 years), high (more than 11 years)), level of mathematics qualifications (specialist qualification No = 0, Yes = 1), and experience of professional learning development (No = 0, Yes = 1). School socioeconomic level and

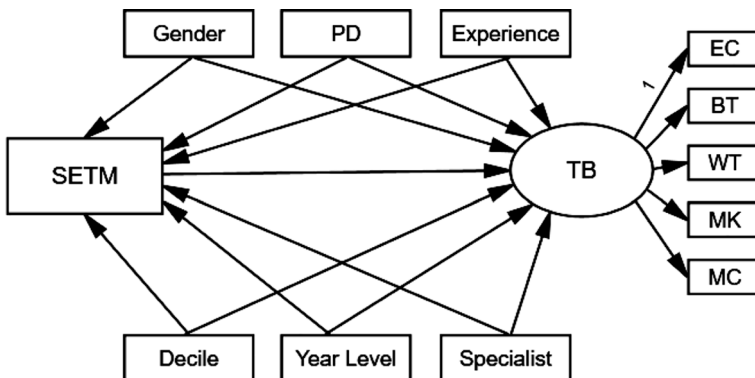


Fig. 2 Multilevel SEM model of the study

Table 3 Confirmatory factor analysis of scales

Scale	χ^2	<i>p</i>	RMSEA (90% CI)	CFI	WRMR
Teachers' self-efficacy in teaching mathematics (SETM)	17.187	0.016	0.067 (0.027, 0.107)	0.994	0.540
An ethic of care (EC)	4.210	0.122	0.058 (0.000, 0.137)	0.966	0.448
Building on students' thinking (BT)	3.840	0.147	0.053 (0.000, 0.133)	0.978	0.435
Worthwhile mathematical tasks (WT)	4.166	0.526	0.000 (0.000, 0.070)	10.000	0.341
Making connections (MK)	3.482	0.175	0.048 (0.000, 0.129)	0.990	0.366
Mathematical communication (MC)	8.390	0.136	0.046 (0.000, 0.097)	0.970	0.560

RMSEA root-mean-square error of approximation, *CFI* comparative fit index, *WRMR* weighted root-mean-square residual

Table 4 Standardized factor loadings of pedagogical practices scales

	SETM	EC	BT	WT	MK	MC
Item 1	0.86	0.42	0.45	0.32	0.52	0.37
Item 2	0.78	0.29	0.65	0.40	0.61	0.56
Item 3	0.78	0.76	0.36	0.40	0.50	0.52
Item 4	0.49	0.39	0.60	0.66	0.68	0.73
Item 5	0.60			0.74		0.41
Item 6	0.80					
ICC	0.19	0.15	0.09	0.26	0.10	0.10

teaching experience variables were dummy coded with low decile and less experience (0–5 years) as reference groups, respectively. We hypothesized that SETM positively influences how teachers teach (TB) and selected teacher/school characteristics associated with both SETM and TB.

Results

Confirmatory factor analyses (CFA)

To determine to what extent there was evidence to support the factorial structure of the SETM scale and the pedagogical practices scales, we hypothesized each scale to be unidimensional, which assumed all manifest variables loaded on one single factor. The goodness-of-fit measures are summarized in Table 3.

The overall fit measures indicated that the hypothesized models were sufficient representations of the structures underlying the constructs. Chi-square values were non-significant for all except the SETM scale. However, χ^2/df was less than 3 for this scale. All RMSEA and CFI values were within the commonly accepted standards. Standardized factor loadings (Table 4) ranged from 0.29 to 0.86, and all were significant at $p < 0.001$. (Note: The scales had different numbers of items. The first factor loading for the SETM scale is 0.86. That is the loading for the first item on the SETM scale. The first factor loading for EC scale was 0.42. That is the loading for the first item on the EC scale. Thus, “Item 1”

refers to the first item on each individual scale.) The psychometric properties we provided here indicated that those items group together and could be reported on at the scale level.

A recommended initial step in fitting a multilevel model involves justifying the need for a multilevel approach by reporting the degree of clustering in the outcome measures. The estimated ICCs, as outlined in Table 4, range from 0.09 to 0.19, necessitating the employment of a multilevel analysis. Thus, a multilevel SEM model was employed for further analyses.

Descriptive statistics

After establishing an acceptable fit for the SETM measurement model, we used this scale to understand the New Zealand teachers' SETM overall. To simplify interpretation of the results, we summed the scores of participants across the six items, and then divided by six. Thus, the mean scores can be referenced against the 1–4 response categories of the items. As shown in Table 5, the means for SETM items ranged from 3.28 to 3.51, suggesting that most teachers indicated either “moderately true for me” or “very true for me.” The mean SETM scale score for all teachers was 3.39. Year 8 teachers reported higher ($M=3.45$, $SD=0.44$) self-efficacy in teaching mathematics than Year 4 teachers ($M=3.33$, $SD=0.46$).

Overall, therefore, these 327 New Zealand teachers reported they felt moderately self-efficacious when teaching mathematics. In other words, overall, they were confident about teaching mathematics and felt they had the necessary knowledge and skills to teach the subject to a diverse range of students. They felt able to respond to difficult questions and felt they could provide alternative explanations or examples. They believed that they could gauge how well the students understood their teaching and felt able to motivate students who showed little interest in mathematics.

Multilevel SEM modeling

The estimated multilevel SEM model displayed a satisfactory fit to the data ($\chi^2=82.58$, $df=55$, $RMSEA=0.040$, $CFI=0.95$, $TLI=0.93$, $SRMR=0.04$). The standardized coefficients of the structural parameters that were estimated from the multilevel SEM analysis are presented in Table 6 among the variables we examined.

The multilevel SEM results showed that SETM had a moderate direct influence on TB ($\beta=0.47$, $p<0.001$), meaning that teachers with high SETM engaged in effective teaching pedagogies more often than teachers with low SETM. After controlling for all the other

Table 5 Means and standard deviations for SETM items

	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
SETM 1	2	4	3.51	0.56	−0.60	−0.66
SETM 2	1	4	3.45	0.66	−1.06	1.02
SETM 3	1	4	3.28	0.66	−0.45	−0.46
SETM 4	2	4	3.32	0.59	−0.23	−0.63
SETM 5	1	4	3.42	0.67	−0.91	0.46
SETM 6	1	4	3.38	0.62	−0.54	−0.20

Table 6 Standardised Coefficients of Structural Parameters from Multilevel SEM Analysis

	SETM			TB		
	Estimate	SE	<i>p</i>	Estimate	SE	<i>p</i>
Level 1: Teachers						
SETM				0.47	0.07	0.00
Year level	0.09	0.06	0.15	-0.05	0.07	0.49
Gender	0.02	0.06	0.72	0.12	0.05	0.03
Experience						
High versus low	0.34	0.08	0.00	-0.03	0.09	0.77
Mid versus low	0.17	0.08	0.04	0.05	0.18	0.59
Specialist	0.23	0.04	0.00	-0.09	0.07	0.17
PD	0.10	0.05	0.06	0.16	0.06	0.01
Level 2: Schools						
SES (Decile)						
High versus low	0.09	0.27	0.75	-0.26	0.18	0.14
Mid versus low	0.33	0.30	0.28	0.02	0.20	0.93

Significant coefficients ($p < 0.05$) are highlighted in bold

variables in the model, female teachers had higher ($\beta = 0.12$, $p = 0.03$) TB scores than their male peers. Teachers were asked to indicate if they had any opportunities in the last 12 months for professional learning and development (PD) focused on mathematics. We found that teachers who received PD had higher TB scores ($\beta = 0.16$, $p = 0.01$) than those who had not received PD. The influence of year level and school socioeconomic status on TB were not significant.

Teachers were asked if they had specialist qualifications in mathematics (such as a mathematics major in their teacher education degree and/or mathematics-focused university papers). In relation to SETM, findings revealed that more experienced teachers (high experience: $\beta = 0.34$, $p < 0.001$, mid experience: $\beta = 0.17$, $p = 0.04$) and teachers with specialist mathematics qualification ($\beta = 0.23$, $p < 0.001$) exhibited higher levels of SETM than low experienced teachers and teachers with no specialist mathematics qualification. Teachers did not differ significantly on their SETM scores with respect to their year level, gender, school SES and whether they had PD in the last 12 months.

Our results also indicated that SETM mediated teacher experience-TB and specialist mathematics qualification-TB relationships. This means that as teachers get specialist mathematics qualifications or their work experience increases, so does their SETM, which leads to an increase in their engagement in effective teaching behavior.

Discussion

In this study, we used data from New Zealand's NMSSA project, which included questionnaire data from a nationally representative sample of 327 teachers in Year 4 and Year 8 in New Zealand. We sought to understand how teachers' SETM is related to their teacher/school characteristics and then to their pedagogical practices. Specifically, we sought to examine the relationship between SETM and characteristics of teachers and the schools in which they taught, and the relationship between SETM and the self-report of effective teaching practices.

Some of our non-significant findings were most encouraging. For example, we found no significant differences in SETM between male and female teachers, and no differences in relationship to the socioeconomic characteristics of the schools in which the teachers taught. As noted in the literature review, there is not extensive nor conclusive work in either of these areas to compare our findings to. Thus, we note that we did not find gender or SES effects here, but caution that the context of New Zealand education should be considered when trying to generalize these findings. We also note that experience was not related to teaching behaviors.

We found that Year 8 mathematics teachers had significantly higher SETM scores than their Year 4 colleagues when year level was used as the only predictor in the model. However, after controlling for other variables included in the model this difference lessened and became non-significant. This is an interesting finding in the light of prior research that has identified an inverse relationship between teachers' self-efficacy and year level of students (Durksen et al., 2017; Klassen & Chiu, 2010; Wolters & Daugherty, 2007). In the New Zealand school context, Year 8 is the final year of primary education and is taught by primary-level trained teachers that are drawn from the same population as their colleagues working in Year 4.

More experienced teachers at both year levels had higher SETM scores than their less experienced peers. This is consistent with findings from Klassen and Chiu (2010), who reported a rise in TSE over the first two decades of teaching. Interestingly, Klassen and Chiu reported a gradual decline in TSE beyond 23 years of service, whereas our results show continued growth, albeit somewhat slowed, of SETM, a domain-specific form of teacher self-efficacy (TSE), for teachers with more than 20 years-service. Unfortunately, we were not able to further separate the teachers with over 23 years-service to make a direct comparison with their findings. Nevertheless, given the importance of robust TSE, this finding is a reminder to school leaders and ministries of education that benefits come with experienced teachers. Furthermore, the finding suggests that in-service teacher education support should be available to scaffold the development SETM in early career teachers.

Teachers with specialist mathematics qualifications exhibited significantly higher SETM scores at both year levels than teachers without these qualifications. Yet, this group represented only 13.9% of Year 4 teachers and 23.3% of Year 8 teachers. This may be of interest to those charged with recruiting and selecting prospective teachers. Highly self-efficacious mathematics teachers were more likely to have recently received professional learning experiences and were more likely to rate the level of professional support they received in their work as "good" or "excellent" than those with mid or low SETM. These results highlight the relationship between SETM and professional learning and are in keeping with those from other teachers; self-efficacy studies, such as that of Durksen and colleagues (2017). Consonant with these authors, we believe these results are best explained in terms of Bandura's (1997) notion of reciprocal determinism, in that professional learning involves the interplay of personal, behavioral and environmental factors.

Regarding the question of the relationship between SETM and reported teaching practices, there was a moderately significant relationship between SETM and TB. Teachers who were highly self-efficacious reported using pedagogical practices that are consistent with Building on Students' Thinking (the difference was most prominent for this scale), providing Worthwhile Mathematical Tasks, supporting students in Making Connections, and encouraging Mathematics Communication. In other words, teachers who were self-efficacious in their mathematics teaching reported more frequent use of pedagogical practices known to be effective in the mathematics classroom than their lower efficacy colleagues.

Earlier, we noted the difficulty of attributing the direction of influence in correlational studies of teachers' self-efficacy. This is true of this study as well, however we found that high SETM and good teaching behaviors go hand in hand. Attributing directionality exclusively in one direction or the other is probably unrealistic, and unnecessary. It seems to us more reasonable to imagine that there is a degree of reciprocal relationship here. We modeled the relationship with SETM influencing teaching behaviors and found good fit to that but we believe this relationship—between one's confidence in ability to teach and how one conducts teaching—to be worthy of further development and research. As mentioned earlier, we chose the directionality of SETM to TB on a logical argument: that SETM influences instructional choices, but it is the mastery experience of a successful classroom using those instructional practices that leads to increased SETM, not taking the choices themselves. We also found that SETM is related to mathematics qualifications and teaching experience. Thus, it makes good sense to encourage those with expertise and interest in mathematics to become primary teachers. This finding is likely to support discussions about the selection of new preservice teachers and the nature of their programs of preparation.

We also found that SETM is related to mathematics qualifications and teaching experience. Thus, it makes good sense to encourage those with expertise and interest in mathematics to become primary teachers. This finding is likely to support discussions about the selection of new preservice teachers and the nature of their programs of preparation.

There are limitations to our study which need to be kept in mind when considering these results. NMSSA used a two-stage stratified sampling design to seek a nationally representative samples of primary school students, and the teachers of these students were then invited to complete make up a sample of teachers. Although the response rate was very strong, still some schools declined the initial invitation to participate and other schools were not included, such as private schools and correspondence schools. Thus, the data were not a true random sample, but we are confident that the sample was representative of teachers at these levels in New Zealand, and therefore, we believe, an exceptionally strong data set for examining these questions.

These results are reliant on self-reports of teachers' pedagogical practices. As with many affective variables, there is a complex relationship between teachers' espoused beliefs and their practice (Beswick, 2005). Thus, there may be some tendency for teachers to report what they believe the answer should be, and not necessarily consistent with the pedagogy they enact within the classroom. Furthermore, the SETM items are not worded as advised by Bandura (2006), but to provide information to New Zealand's Ministry of Education. Notwithstanding these limitations, we believe that the findings presented here are representative of teachers at these year levels in New Zealand, and that for the most part, teachers gave good faith responses to our questions. These results are based on a major research initiative in New Zealand that has been running for over 30 years; teacher responses were anonymous; and the NMSSA program is held in high regard among New Zealand educators. So, while some caution is appropriate, we believe the findings presented here are a strong contribution to the literature in teacher self-efficacy.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the

material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allan, R. (2012). National monitoring study of student achievement: Wanangatia te Putanga Tauira. *Set: Research Information for Teachers*. <https://doi.org/10.18296/set.0371>
- Anthony, G., & Walshaw, M. (2009a). *Effective pedagogy in mathematics* (Vol. 19). International Academy of Education.
- Anthony, G., & Walshaw, M. (2009b). Characteristics of effective teaching of mathematics: A view from the west. *Journal of Mathematics Education*, 2(2), 147–164.
- Attard, C., Ingram, N., Forgasz, H., Leder, G., & Grootenboer, P. (2016). Mathematics education and the affective domain. In K. Makar, S. Dole, J. Visnovska, M. Goos, A. Bennison, & K. Fry (Eds.), *Research in mathematics education in Australasia 2012–2015* (pp. 73–96). The Netherlands: Springer.
- Asil, M. (2017). A school-based measure of culturally responsive practices. *Frontiers in Education*, 2(17), 1–8.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117–148.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Information Age.
- Beatson, N. J., Berg, D. A., & Smith, J. K. (2018). The impact of mastery feedback on undergraduate students' self-efficacy beliefs. *Studies in Educational Evaluation*, 59, 58–66.
- Berg, D. A. G., (2022). Over 40 years of teacher self-efficacy research, theory, measurement, and application. In A. O'Donnell, N. Barnes, & J. Reeves (Eds.), *Oxford handbook of educational psychology*. Oxford University Press.
- Berg, D. A. G., Skaalvik, E. M., Asil, M., Hill, M. F., Uthus, M., Tangen, T. N., & Smith, J. K. (2023). Teacher self-efficacy and reasons for choosing initial teacher education programmes in Norway and New Zealand. *Teaching and Teacher Education*, 125, 104041.
- Beswick, K. (2005). The beliefs/practice connection in broadly defined contexts. *Mathematics Education Research Journal*, 17(2), 39–68.
- Bjerke, A. H., & Xenofontos, C. (2023). Teachers' self-efficacy in teaching mathematics: Tracing possible changes from teacher education to professional practice. *Teachers and Teaching*, 30, 1–15.
- Blotnicky, K. A., Franz-Odenaal, T., French, F., & Joy, P. (2018). A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *International Journal of STEM Education*, 5(1), 1–15.
- Brouwers, A., Evers, W., & Tomic, W. (2001). Self-efficacy in eliciting social support and burnout among secondary-school teachers. *Journal of Applied Social Psychology*, 31(7), 1474–1491.
- Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, 16(2), 239–253.
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. Guildford.
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research*, 21(2), 230–258.
- Carney, M. B., Brendefur, J. L., Thiede, K., Hughes, G., & Sutton, J. (2016). Statewide mathematics professional development: Teacher knowledge, self-efficacy, and beliefs. *Educational Policy*, 30(4), 539–572.
- Cayhill, R., & Zhao. (2020). *Progress and achievement and the context of mathematics and statistics learning in New Zealand*. <https://assets.education.govt.nz/public/Literacy-and-Maths-strategy-development-in-2021/Progress-and-achievement-and-the-context-of-mathematics-and-statistics-learning-in-New-Zealand.pdf>
- Chatzistamatiou, M., Dermizaki, I., & Bagiatis, V. (2014). Self-regulatory teaching in mathematics: Relations to teachers' motivation, affect and professional commitment. *European Journal of Psychology of Education*, 29(2), 295–310.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.

- Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social-emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology, 104*(4), 1189.
- Cousins, J. B., & Walker, C. A. (2000). Predictors of educators' valuing of systematic inquiry in schools. *Canadian Journal of Program Evaluation, 15*(3), 25–52.
- Darr, C., Joyce, C., & Burgon, J. (2017). The national monitoring study of student achievement: Wānangatia te Putanga Tauira. *Journal Issue*. <https://doi.org/10.18296/set.0083>
- Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education, 24*(3), 751–766.
- Durksen, T. L., Klassen, R. M., & Daniels, L. M. (2017). Motivation and collaboration: The keys to a developmental framework for teachers' professional learning. *Teaching and Teacher Education, 67*, 53–66.
- Eynde, P. O. T., De Corte, E., & Verschaffel, L. (2006). "Accepting emotional complexity": A socio-constructivist perspective on the role of emotions in the mathematics classroom. *Educational Studies in Mathematics, 63*, 193–207.
- Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling, 12*(3), 343–367.
- Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research, 42*(3), 509–529.
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*, 466–491.
- Gabriele, A. J., & Joram, E. (2007). Teachers' reflections on their reform-based teaching in mathematics: Implications for the development of teacher self-efficacy. *Action in Teacher Education, 29*(3), 60–74.
- Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods, 19*(1), 72–91.
- Ghaith, G. Y., & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 13*(4), 451–458.
- Grootenboer, P., & Marshman, M. (2016). *Mathematics, affect and learning: Middle school students' beliefs and attitudes about mathematics education*. Springer Science+Business Media.
- Guo, Y., Connor, C. M., Yang, Y., Roehrig, A. D., & Morrison, F. J. (2012). The effects of teacher qualification, teacher self-efficacy, and classroom practices on fifth graders' literacy outcomes. *The Elementary School Journal, 113*(1), 3–24.
- Guo, Y., Piasta, S. B., Justice, L. M., & Kaderavek, J. N. (2010). Relations among preschool teachers' self-efficacy, classroom quality, and children's language and literacy gains. *Teaching and Teacher Education, 26*(4), 1094–1103.
- Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice Hall.
- Hardré, P. L. (2011). Motivation for math in rural schools: Student and teacher perspectives. *Mathematics Education Research Journal, 23*(2), 213–233.
- Hilbert, S., Bruckmaier, G., Binder, K., Krauss, S., & Bühner, M. (2019). Prediction of elementary grades by cognitive abilities. *European Journal of Psychology of Education, 34*, 665–683.
- Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of Educational Psychology, 105*(3), 774.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to under parameterized model misspecification. *Psychological Methods, 3*, 424–453.
- Ingram, N. (2017). We are still learning to integrate affect (and mathematics) into our research. In A. Downton, S. Livy, & J. Hall (Eds.), *40 years on: We are still learning! (Proceedings of the 40th Annual Conference of the Mathematics)*.
- Ingram, N. (2020). Growing mathematics learners. *New Zealand Annual Review of Education, 26*, 190–196.
- Ingram, N., Hatisaru, V., Grootenboer, P., & Beswick, K. (2020). Researching the affective domain in mathematics education. *Research in Mathematics Education in Australasia 2016–2019*, 147–175.
- Jacobs, V. R., & Spangler, D. A. (2017). Research on core practices in K-12 mathematics teaching. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 766–792). National Council of Teachers of Mathematics.
- Khine, M. S. (2017). Spatial cognition: Key to STEM success. In M. S. Khine (Ed.), *Visual-spatial ability in STEM education*. Springer International Publishing.
- Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102*(3), 741–756.
- Klassen, R. M., Tze, V. M., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: Signs of progress or unfulfilled promise? *Educational Psychology Review, 23*(1), 21–43.

- Klassen, R., Wilson, E., Siu, A. F., Hannok, W., Wong, M. W., Wongsri, N., Sonthisap, P., Pibulchol, C., Buranachaitavee, Y., & Jansem, A. (2013). Preservice teachers' work stress, self-efficacy, and occupational commitment in four countries. *European Journal of Psychology of Education, 28*(4), 1289–1309.
- Lee, J., & Shute, V. J. (2010). Personal and social-contextual factors in K–12 academic performance: An integrative perspective on student learning. *Educational Psychologist, 45*(3), 185–202.
- Lomas, L., & Clarke, D. (2016). Changes in teachers' knowledge and beliefs about mathematics and mathematics teaching: A case study. In B. White, M. Chinnappan, & S. Trenholm (Eds.), *Proceedings of the 39th annual conference of the mathematics education research group of Australasia* (pp. 391–398). Adelaide: MERGA.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.
- Malmberg, L. E., Hagger, H., & Webster, S. (2014). Teachers' situation-specific mastery experiences: Teacher, student group and lesson effects. *European Journal of Psychology of Education, 29*(3), 429–451.
- Ministry of Education. (2007). *The New Zealand curriculum*. Learning Media Limited.
- Ministry of Education. (2015). *National monitoring study of student achievement: Mathematics and statistics 2013*. Ministry of Education.
- Murphy, S., & Ingram, N. (2023). A scoping review of research into mathematics classroom practices and affect. *Teaching and Teacher Education, 132*, 104235.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus: Statistical analysis with latent variables: User's guide*. Los Angeles: Muthén & Muthén.
- Perera, H. N., & John, J. E. (2020). Teachers' self-efficacy beliefs for teaching math: Relations with teacher and student outcomes. *Contemporary Educational Psychology, 61*, 101842.
- Philippou, G., & Christou, C. (2002). A study of the mathematics teaching efficacy beliefs of primary teachers. In *Beliefs: A hidden variable in mathematics education?* (pp. 211–231). Dordrecht: Springer Netherlands.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models*. Sage.
- Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education, 65*, 150–167.
- Roberts, J. K., & Henson, R. K. (2001). A confirmatory factor analysis of a new measure of teacher efficacy: Ohio State teacher efficacy scale. In *A paper presented at the meeting of the American Educational Research Association, Seattle*.
- Russo, J., Bobis, J., Downton, A., Hughes, S., Livy, S., McCormick, M., & Sullivan, P. (2020). Elementary teachers' beliefs on the role of struggle in the mathematics classroom. *The Journal of Mathematical Behavior, 58*, 100774.
- Saadati, F., Chandia, E., Cerda, G., & Felmer, P. (2021). Self-efficacy, practices, and their relationships; The impact of a professional development program for mathematics teachers. *Journal of Mathematics Teacher Education, 26*, 103–124.
- Sarac, A., & Aslan-Tutak, F. (2017). The relationship between teacher efficacy, and students' trigonometry self-efficacy and achievement. *International Journal for Mathematics Teaching & Learning, 18*, 66–83.
- Sass, D. A., Schmitt, T. A., & Marsch, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multi-disciplinary Journal, 21*(2), 167–180.
- Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*(3), 611–625.
- Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education, 26*(4), 1059–1069.
- Skaalvik, E. M., & Skaalvik, S. (2017). Still motivated to teach? A study of school context variables, stress and job satisfaction among teachers in senior high school. *Social Psychology of Education, 20*(1), 15–37.
- Skilling, K., Bobis, J., Martin, A., Anderson, J., & Way, J. (2016). What secondary teachers think and do about student engagement in mathematics. *Mathematics Education Research Journal, 28*(4), 545–566.
- Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *The Elementary School Journal, 110*(2), 228–245.
- Tschannen-Moran, M., Woolfolk-Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202–248.
- Wolters, C. A., & Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology, 99*(1), 181.
- Woolfolk Hoy, A., & Davis, H. (2005). Teachers' sense of efficacy and adolescent achievement. *Adolescence and Education, 5*, 117–137.
- Yu, C. Y. (2002). *Evaluating cut-off criteria of model fit indices for latent variable models with binary and continuous outcomes* (Doctoral dissertation, University of California Los Angeles).

Zee, M., & Koomen, M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, 86(4), 981–1015.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.