Teaching and Teacher Education 96 (2020) 103190

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Research paper

Do teachers' years of experience make a difference in the quality of teaching?

Linda J. Graham [a, *], Sonia L.J. White [a], Kathy Cologon [b], Robert C. Pianta [c]

[a] Queensland University of Technology (QUT), Australia
[b] Macquarie University, Australia
[c] University of Virginia, USA

HIGHLIGHTS

- Teaching quality observed for 80 teachers with 0–3, 4–5, and >5 years' experience.
- No evidence of less competence for teachers with 0–3 years' experience.
- Some evidence of decline in teaching quality in teachers with 4–5 years' experience.
- Behaviour management among highest scoring dimensions for all teachers.
- Support and professional learning a priority for all teachers, not just beginners.

ARTICLE INFO

Article history:
Received 17 January 2020
Received in revised form
3 August 2020
Accepted 13 August 2020
Available online 29 August 2020

Keywords:
Initial teacher education
Beginning teachers
Teachers' years of experience
Teaching quality
Classroom observations

ABSTRACT

Extensive reforms have been made to initial teacher education (ITE) to improve "teacher quality" without any evidence to support the claim that beginning teachers are less competent than experienced teachers. This study adds to the evidence base by investigating associations between teachers' years of experience and teaching quality. Results show no evidence of lower teaching quality for beginning teachers (0–3 years' experience), but some evidence of a decline in teaching quality for teachers with 4–5 years experience. Findings suggest that the quality of teaching could be higher overall, and that targeted support and evidence-informed professional learning would benefit all teachers.

1. Introduction

The international comparison of student performance in standardised assessments and subsequent ranking of participating countries over the last two decades has generated significant public and political anxiety, leading to an intense focus on factors believed to influence the academic achievement of school students (Sellar, Thompson, & Rutkowski, 2017). In the 20 years since the OECD's *Programme for International Student Assessment* (PISA) began, attention has shifted from in/equity in student achievement to declines in both relative and absolute performance, leading to increased scrutiny of teaching and "teacher quality" (Scholes et al., 2017). This shift in focus has been propelled by claims that "in excess of 40 percent of the residual variance in measures of student performance (adjusted for students' background and intake characteristics) is at the class/teacher-level" (Ingvarson & Rowe, 2008, p. 7). Increased scrutiny of teachers through the development of value-added models of teacher effectiveness based on student achievement in national standardised assessment systems, the establishment of statutory authorities to develop and oversee teacher professional standards and certification, and the reinstatement of formal inspections, has followed (Baxter, 2013; Sachs, 2016).

Since 2009, edu-tourism—where delegations of politicians, principals, and teachers visit education systems at the top of the international league table at the hope of learning from their success—has swung from the "Finnish miracle" to the "Asian tigers" of

* Corresponding author. The Centre for Inclusive Education, Faculty of Education, Victoria Park Road, Kelvin Grove, QLD, 4059, Australia.
E-mail address: linda.graham@qut.edu.au (L.J. Graham).

https://doi.org/10.1016/j.tate.2020.103190

Singapore, Hong Kong, South Korea and Shanghai (Lingard & Sellar, 2019). Of interest has been the "Maths Mastery" approach to teaching mathematics, which was introduced and tested in England through two randomised controlled trials involving 90 primary schools and 50 secondary schools (Jerrim & Vignoles, 2016). Also making the news headlines are reported differences between systems in terms of classroom noise and disorder with the conclusion being that teachers in lower ranked countries, like Australia (ranked 70th of 77 countries), are poor at behaviour management (Bolton, 2019; Carey, 2019). Blunt analyses of Teaching and Learning International Survey (TALIS) data indicating higher scores for experienced versus novice teachers give the impression that early career teachers are less effective, despite the caveat that beginning teachers are disproportionately allocated to more challenging schools (Organisation for Economic Co-operation and Development, 2019).

Scrutiny has since moved from system-level effects—such as inequities in school funding and the concentration of advantage/disadvantage through the social segregation engineered by competitive school markets—to the effectiveness of teacher preparation (Gale & Parker, 2017; Mayer et al., 2017; McMahon, Forde, & Dickson, 2015). University initial teacher education (ITE) has been criticised for being too theoretical and inadequately preparing teachers for the practical realities of the classroom (Teacher Education Ministerial Advisory Group [TEMAG], 2014). This criticism is particularly acute with regards to behaviour management (Joseph, 2017; O'Neill & Stephenson, 2012), which has been branded as 'ideological' (Bolton, 2019) without any rigorous investigation of what is being taught in ITE, whether this is really where the problem lies, or whether beginning teachers are any less effective at behaviour management than more experienced teachers. The effect of coupling ITE and "teacher quality," however, is that it frames graduate or beginning teachers as "the problem" and this, in turn, frames the development of solutions (Mockler, 2018). In other words, a narrow focus on ITE and the graduates it produces may mean that the true nature and breadth of the problems impacting school education remain undetected and unresolved, while others are magnified beyond their actual or practical significance.

A deficit framing of ITE is particularly prominent in Australia (Churchward & Willis, 2018), where university teacher education replaced Teachers College in the 1990s and where most practising teachers are now degree qualified. Recent solutions to "the problem" of quality teaching in Australia have therefore focused predominantly on university teacher education and the quality of ITE graduates (Gore, 2016). Australian universities are now under pressure to produce "classroom ready" graduates with greater prescription of their course content through formal accreditation processes (Churchward & Willis, 2018; Gale & Parker, 2017). In 2016, for example, an external assessment of the literacy and numeracy capability of teacher education students was introduced by the Australian government as a condition of graduation to ensure "increased confidence in the skills of graduating teachers" (DET, 2016 as cited in McGraw & Fish, 2018, p. 121). A further graduation hurdle for teacher education students, introduced in 2018 and linked to the accreditation of university teacher education programs (Churchward & Willis, 2019), is successful completion of the Teacher Performance Assessment (TPA) "which uses evidence of classroom readiness determined by impact on student learning and is associated with the Australian Professional Standards for Teachers" (p. 252).

Further, and as debate continues to rage over the sufficiency or otherwise of university ITE entry scores (Morgan & Aspland, 2018), the government responsible for Australia's largest education system in the state of New South Wales has mandated that graduates will need to demonstrate "superior intelligence" in a psychometric test and maintain a credit grade point average in order to register as a teacher upon graduation (NSW Government Department of Customer Service, 2019). Other changes—including the increase in and privileging of professional experience in ITE over theory (Gale & Parker, 2017) as well as the "stage" approach in the Australian Professional Standards for Teaching, which ranges from Graduate, Proficient, Highly Accomplished, and Lead—are predicated not only on the assumption that there is a difference between graduate and proficient, but that there is a positive relationship between experience and teaching quality. All these reforms, however, have occurred without the provision of any empirical evidence to support the claim that beginning teachers are less competent than teachers with more years' experience, especially in behaviour management. This may be due to the limited availability of empirical evidence and because the evidence that does exist is mixed.

1.1. Evidence of an association between experience and teaching quality

At a political level and within public discourse, there is a tendency to presume a straightforward linear relationship between teachers' years of experience and the quality of teaching (Brandenburg, McDonough, Burke, & White, 2016). By contrast, research demonstrates a complex relationship between a range of factors that is non-linear and cyclical, whereby experience is one of many factors influencing the quality of teaching (Brandenburg et al., 2016; Klassen & Chiu, 2010). Attempts to research this complex relationship has resulted in mixed findings (Chingos & Peterson, 2011; Rivkin, Hanushek, & Kain, 2005), which have been complicated by the range and incommensurability of measures used in research, as well as inconsistency in the conceptualisation of teaching quality and the nomenclature used to describe categories of experience. For example, the terms "graduate," "beginning," and "early career" are used interchangeably to describe teachers with any combination of experience of up to five years (Gallant & Riley, 2014; Mockler, 2018; Sullivan, Johnson, & Simons, 2019). Adding further confusion, are studies that categorise teachers as "novice/expert" or "beginner/experienced" without specifying the number of years that constitute these experience categories (Palmer, Stough, Burdenski, & Gonzales, 2005). Even in studies where experience categories are defined, there is little consistency. For example, Araujo, Carneiro, Cruz-Aguayo, and Schady (2016) define less experienced or "rookie" teachers as those with 0–3 years' experience and "experienced" teachers as those with more than three years' experience. The groups in the Araujo et al. study, however, are incommensurable with earlier research using the same observational measure of quality teaching, conducted by Cortina, Miller, McKenzie, and Epstein (2015). These researchers describe "experienced" teachers as those who mentored student teachers, and "novice" teachers as first year graduates who taught the same class as their mentor teacher.

There are also differences in the conceptualisation of teaching quality. In some studies, quality is a euphemism for teacher effectiveness, which is predominantly measured indirectly using student test scores. Higher test scores are presumed to indicate the presence of quality teaching; however, these analyses are driven by an "outcome led conception of teaching" (Flores, 2019), tend not to define quality, are agnostic on the problem of test score manipulation through teaching to the test and/or excluding students with learning difficulties (Lauen & Gaddis, 2016), and—by using academic achievement as a proxy measure—ignore domain-specific elements of quality teaching that contribute to students' social and emotional development (Hamre, Hatfield, Pianta, & Jamil,

2014). Other studies use direct measures that articulate a specified range of quality teaching practices and then observe for them. In these studies, teaching quality is conceptualised as a multidimensional construct that is theoretically derived from empirical research identifying a range of teaching practices that make a positive contribution to students' emotional, behavioural, and academic outcomes (Pianta and Hamre, 2009).

1.1.1. Evidence from studies using indirect measures

The most common indirect measure of teaching quality is student performance in standardised assessments and findings from this research are mixed. There is some evidence that teachers' years of experience have little or no impact on student outcomes (Rivkin et al., 2005; Stuhlman & Pianta, 2009), however, there is also evidence to the contrary (Araujo et al., 2016; Podolsky, Kini, & Darling-Hammond, 2019). Again though, this evidence is not clear cut, suggesting that there is some effect for experience but that this is limited and not cumulative. For example, some research across national and international contexts provides evidence of an early effect for experience whereby beginning teachers quickly improve, but this association declines after their initial adjustment to the field (Chingos & Peterson, 2011; Klassen & Chiu, 2010; Rivkin et al., 2005; Rockoff, 2004). Similarly, in a large North American study drawing on data made available by the Florida Department of Education involving 84,031 teachers across Grades 4 to 8, Chingos and Peterson (2011) found that teachers generally become more effective in the areas of mathematics and reading after their initial transition into the field (particularly after the first year), but the return to experience plateaus or declines after four to five years. In another North American study using panel data on 300 teachers and standardised test scores for 10,000 students, Rockoff (2004) found a small but positive impact of teacher experience in the first two years of teaching for math computation, and in the first 10 years of experience for vocabulary and reading comprehension. In a similar study in North America involving data from more than half a million students in over 3000 schools across Grades 3 to 7, Rivkin et al. (2005) found a positive effect for teaching experience on student outcomes in mathematics and reading for teachers in their first year of teaching. However, these effects did not continue beyond the initial period of transition into the profession.

The evidence for associations between teachers' years of experience and teaching quality from studies using indirect measures (e.g., attitudes, student performance) is by and large mixed and does not suggest a clear set of conclusions regarding the importance of experience as a predictor of quality. The literature also has addressed questions concerning the returns to teacher experience by examining more direct indicators of teaching quality, specifically teachers' observed classroom behaviours in areas of classroom management, social support for students, and instruction (Reyes, Brackett, Rivers, White, & Salovey, 2012; Rucinski, Brown, & Downer, 2018). Given the oft-cited reports that novice teachers describe challenges in classroom management (Voss, Wagner, Klusmann, Trautwein, & Kunter, 2017) and research indicating that increasing years of experience can lead to improved performance in this domain (Ladd & Sorensen, 2017), it stands to reason that examining teachers' actual classroom interactions could be useful in attempts to understand the links between experience and performance. Moreover, given that interactions with students are the most direct and proximal indicator of teachers' influence on students, observations of these interactions are an important indicator of teaching quality.

1.1.2. Evidence from studies using direct observation

Only a small number of studies have investigated the link between teaching experience and teaching quality using direct observation. For example, in North America, Stuhlman and Pianta (2009) investigated the relationship between years of teaching experience and teaching quality in 820 Grade 1 classrooms from 700 schools across 32 states using the Classroom Observation System for First Grade (COS-1; NICHD ECCRN, cited in Stuhlman & Pianta, 2009); an early prototype of the Classroom Assessment Scoring System (CLASS). Teachers' years of experience in this study ranged from less than one to 41 years. Four classroom types were identified: 1) positive emotional climate/lower academic demand, 2) high overall quality, 3) mediocre quality, and 4) low overall quality. No significant differences between the four classroom types were found in relation to teachers' years of experience.

In another North American study, Schachter, Spear, Piasta, Justice, and Logan (2016) investigated the effect of teaching experience on language and literacy instruction with 222 early childhood educators who had zero to 36 years teaching experience. This team used the Individualizing Student Instruction (ISI) classroom observation system (Connor, Piasta et al. cited in Schachter et al., 2016), which was adapted to focus on the language and literacy instruction provided in early childhood education settings. Teacher years of experience was found to be negatively associated with the amount of instruction, with less experienced teachers achieving significantly higher scores. A positive relationship between quantity of instruction, and teaching quality/student learning outcomes was assumed and not investigated in this study.

An effect for teachers' years of experience on student learning outcomes was found in a large study funded by the Inter-American Development Bank involving 24,000 Kindergarten children in Ecuador (Araujo et al., 2016). Teaching quality was assessed using the Classroom Assessment Scoring System (CLASS; La Paro, Pianta, Stuhlman, & Hamre, 2002), in conjunction with assessments of children's math, language, and executive function, conducted at the beginning and end of the Kindergarten year. Participating teachers were grouped into two experience categories: (1) inexperienced or "rookie" teachers with three years' experience or less, and (2) "experienced" teachers, conceptualised as those with more than three years teaching experience. The authors report that Kindergarten children taught by "rookie" teachers achieved test scores that were 0.17 of a standard deviation lower than experienced teachers. However, the study did not compare the CLASS scores of these two groups and there are anomalies in the report. For example, it appears the authors tried to optimise the variance explained in learning outcomes using CLASS and experience. Regressions were run with different definitions of "rookies" (0 years, 0–1 years, 0–2 years … 0-N years) and the maximum variance was 0–3 years of experience, which represented a small six percent of the sample. In their footnote, the authors report that "the returns to experience in Ecuador rise sharply in the first three years, and then flatten out" (p. 11). An alternative interpretation of the iterative regressions, using years of experience, could be that there were some extreme CLASS scores for teachers with three years of experience, driving the finding with a small sample of "rookie" teachers. Analyses investigating teacher transience also found that "rookie" teachers were more likely to move schools, however, no analyses were conducted to assess the potential contribution of transience to lower test scores achieved by the students of those teachers. Mean CLASS scores were also significantly lower for teachers in the Ecuadorian study (3.7) than the cited US sample (4.5) (Araujo et al., 2016), suggesting that findings from this study should be interpreted with caution.

Cortina et al. (2015) compared observation ratings, also using the CLASS, to the eye movement data of 24 teachers (12 experienced or "supervising" teachers and 12 "novices" with less than one years' experience), teaching the same classes in Grades 2 to 8 in North America. The participating teachers wore mobile eye

tracking glasses, and their eye movement data indicated the distribution of visual attention across the class. Cortina et al. (2015) conceptualised teacher quality in classroom monitoring as being equal distribution across all students, rather than focused attention on just a few students. According to this criterion, the experienced teachers scored better in classroom monitoring when compared to the novice teachers, who typically had more focused attention on individual students rather than equal distribution across the whole class. However, when correlating the eye movement patterns against CLASS observation ratings, a more complex relationship was evident, particularly for feedback quality. Novice teachers who had high feedback quality on CLASS ratings also demonstrated greater focused attention on individual students. The novice teachers who had low feedback quality had a more even distribution of visual attention and were considered to have good classroom monitoring, whereas the experienced teachers maintained their evenly distributed classroom visual monitoring irrespective of high or low feedback quality rating. This finding indicates the complexity of quality teaching and the individual differences in approaching feedback. For example, when beginning as a teacher and providing high feedback quality, there may be a cost to overall classroom monitoring—as measured by eye movements—but this does not necessarily mean a reduction in overall teaching quality.

1.2. Summary of evidence across studies

Across studies there is very limited support for the claim that beginning teachers are less competent than teachers with more years of experience. The extant research provides mixed evidence, with some correlations between teaching quality and teacher experience, however these might be conceptualised. However, these findings are confined to a subset of areas of impact and windows of time. Other research provides no evidence that experience makes a difference. Findings from the four studies using direct measures suggest that more experience does not necessarily result in higher quality classroom environments, however, research using direct measures is costly and is therefore rare. The present study aims to contribute to the literature by drawing on standardised classroom observations of 80 primary school classrooms (K-6) and data on years of experience from 80 classroom teachers to investigate whether there is any association between experience and the quality of classroom processes that are important for children's academic, social, and behavioural outcomes (Hamre, Pianta, Mashburn, & Downer, 2007; Pianta & Hamre, 2009).

2. Methodology and methods

2.1. Participants

This study analyses standardised classroom observation data from 18 Preparatory Year teachers, 28 Grade 1 teachers, 20 Grade 2 teachers, and 14 Grade 3 teachers ($n = 80$).[1] Participating teachers were drawn from seven government primary schools in Queensland, Australia, with enrolments ranging from approximately 210–760 children. In 2014, the schools had Index of Community Socio-Economic Advantage (ICSEA) values between 884 and 977; all below the national mean of 1000.[2] The teachers in the 80 classrooms observed were all degree qualified with an average of 9.23 years ($SD = 9.27$) teaching experience. Teachers' years of experience ranged from three weeks through to 38 years. The study commenced in 2014 upon receipt of Queensland University of Technology Human Research Ethics Committee and Queensland Department of Education approvals.

2.2. Data collection

2.2.1. Classroom observations

Classroom interactions were observed annually using the Classroom Assessment Scoring System (CLASS), a standardised observational tool designed to capture and evaluate classroom quality from pre-Kindergarten through to 12th Grade (Hamre et al., 2007). It is used in 30-min observation cycles (20 min observation plus 10 min scoring) with a minimum of four observational cycles per classroom. The measure has demonstrated adequate criterion and predictive validity in studies involving more than 4000 classrooms in the United States (Hamre et al., 2007), mainly due to reliability and objectivity sustained by the mandatory training of observers (Cortina et al., 2015). In this study, classroom observations were conducted using the CLASS K-3 measure, incorporating Prep to Grade 3. Each of the 80 classrooms was observed for a minimum of 2 h (at least four cycles) by certified observers and the first author, a certified CLASS K-3 trainer. In each year of the study, the first 10 observation cycles were conducted by all observers and moderated to ensure consistency. All observations were completed within a two-month period in the third term of each school year. In the CLASS K-3, classroom interactions (teacher-student, student-teacher, student-student) are grouped into three domains and 10 dimensions (Table 1).

For each of the 10 dimensions, ratings based on the observed presence or absence of indicators of quality teaching and learning are made on a 7-point scale, with 1–2 representing scores in the low range, 3–5 representing scores in the mid-range, and 6–7 representing scores in the high range. Mean scores are calculated for the 10 dimensions and then aggregated to determine mean scores for each of the three domains.

The CLASS is used annually by Head Start in the United States to measure classroom quality in grantee programs. In 2016, the domain mean scores for 319 Head Start grantees were 6.00 for Emotional Support, 5.73 for Classroom Organisation, and 2.83 for Instructional Support (Head Start Early Childhood Learning & Knowledge Center, 2016). Large scale studies have shown that mean scores tend to be higher in Emotional Support and Classroom Organisation, than Instructional Support (Pianta & Hamre, 2009). This same research indicates that a threshold of 5.00 and above is needed in both the Emotional Support and Classroom Organisation domains "to promote positive social development and reduce problem behaviours," while a threshold of 3.00 and above in the Instructional Support domain is necessary "to achieve greater gains in early academic and language skills" (The National Center on Quality Teaching and Learning, 2011).

2.2.2. Teacher years of experience

Participating teachers were asked how long they had been teaching and these data were recorded in the main statistical dataset. We conducted analyses in two stages guided by the research literature investigating teaching quality and returns on experience, which suggests two critical timepoints of increase,

[1] The teachers reported here are those who participated in both classroom observations and individual interviews. In the event of a teacher participating multiple times in the longitudinal study, we used only the first year of data collected for that teacher, to avoid replicating individual scores that may skew analyses, either positively or negatively. In this way we have prioritised a range of teachers, rather than repeating individual teacher scores.

[2] The ICSEA values are derived from child-level data, which includes parent occupation and education, and community-level data, which includes remoteness and percent Indigenous enrolment. With a national mean of 1000 (±100), ICSEA values range from approximately 500 (schools with children with extremely educationally disadvantaged backgrounds) to approximately 1300 (schools with children with very educationally advantaged backgrounds).

Table 1
The classroom assessment scoring system (CLASS) domains and dimensions.

Emotional Support	Positive Climate	Polite language and listening are aspects of respect within a class. Teachers demonstrate positive regard by smiling, making eye contact, and using children's names to acknowledge and encourage effort.
	Teacher Sensitivity	Addressing both academic and emotional needs, teacher sensitivity measures the teacher's moment-to-moment response to students cues for support and help.
	Negative Climate[a]	Presence of negative attitudes and behaviours that can include sarcasm, criticism, humiliation, hostility, harsh tone, threats, shouting and physical aggression.
	Regard for Student Perspectives	Teachers' interactions with students and classroom activities should place an emphasis on students' interests, points of view and independence to encourage student responsibility and autonomy.
Classroom Organisation	**Behaviour Management**	Proactive strategies are implemented effectively together with clear expectations, consistency, and clarity of rules with careful monitoring to keep students on task and to anticipate or redirect and address misbehaviour.
	Productivity	Efficient management of instructional time, materials and transitions through prior organisation, appropriate pacing, clear instructions, and use of clear and consistent routines.
	Instructional Learning Formats	The "hooks" teachers use to engage students, such as a variety of materials and modalities, which allow students to access and attend to the instructional opportunities given.
Instructional Support	**Quality of Feedback**	Teachers drive learning by intelligently responding to students' questions, comments and actions, and advance student learning with follow-up questions.
	Concept Development	Measures the teachers' use of instructional discussions and activities to promote students' higher-order thinking skills and cognition, as well as the teacher's focus on understanding rather than on rote instruction.
	Language Modelling	Where teachers intentionally model more advanced, more complex language for students by, for example, introducing, explaining and using new words or synonyms.

[a] Negative climate is reverse scored.

followed by plateau or decline: 0–3 years and 3–5 years, respectively.

In the first stage of analysis, we used two groups to categorise participating teachers' experience (beginning: 0–3 years; experienced: more than 3 years). There were 25 teachers in the 0–3 group, and 55 teachers in the more than 3 years group.

In the second stage of analysis, we used three groupings: beginning: 0–3 years; transitioning: 4–5 years; and experienced: more than 5 years. By distinguishing 0–3 years and 4–5 years, our second analysis reflected two important stages within the 0–5 year period. The 0–3 year category more accurately reflects beginning or novice teacher status than would a 0–5 year period and takes into account research that has identified (i) initial gains for experience within the first few years post-graduation, and (ii) a plateau or decline in teacher effectiveness after three years (Chingos & Peterson, 2011; Klassen & Chiu, 2010; Rivkin et al., 2005; Rockoff, 2004). The 4–5 years category captures a transitional period in which this plateau/decline may commence, and when the majority of early career teacher attrition is said to occur (Chingos & Peterson, 2011; Rivkin et al., 2005). There were 25 teachers in the 0–3 group, 11 teachers in the 4–5 group, and 44 teachers in the more than 5 years group.

2.3. Data analysis

All statistical analyses were completed using SPSS (Ver 25.0, www.ibm.com/spss). Descriptive statistics were used to describe CLASS domains (Emotional Support, Classroom Organisation, and Instructional Support) and subsequently the dimensions that comprise each domain. Univariate analysis of covariance was conducted to examine the differences between the CLASS domain scores for the teacher experience groupings. To adjust for the potential impact of the grade level taught on CLASS scores, grade level taught was included as a covariate in all analyses. A subsequent univariate analysis of covariance examined the effects at a dimension level. These analyses we first conducted with two groups (beginning: 0–3 years; experienced: more than 3 years), then three groups (beginning: 0–3 years; transitioning: 4–5 years; and experienced: more than 5 years). For analysis with the three groups, pairwise comparisons used Bonferroni correction for multiple comparisons.

3. Results

Overall, the 80 participating teachers scored highest in the domains of Emotional Support (M = 5.24, SD = 0.94, Range = 2.13–6.81) and Classroom Organisation (M = 4.90, SD = 1.24, Range = 1.83–6.75), and lowest in Instructional Support (M = 3.60, SD = 1.11, Range = 1.42–5.92). The distribution of CLASS scores is represented in Fig. 1.

3.1. Beginning teachers (0–3 years) and experienced teachers (+3 years)

In the ANCOVA models adjusting for grade level taught, there were no significant differences between the two teacher experience categories (0–3 years and more than 3 years) for any of the three domains (Emotional Support, Classroom Organisation, and Instructional Support) (ps > .05; Table 2). Similarly, when investigating the individual CLASS dimensions, there were no significant differences between the beginning and experienced teacher categories (ps > .05; Fig. 2).

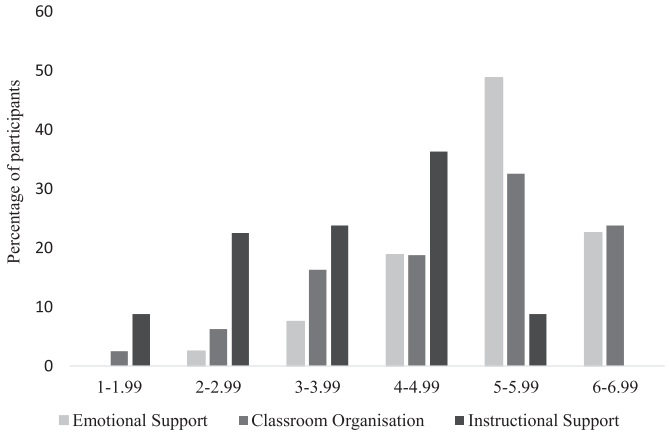

Fig. 1. Distribution of CLASS Scores for each domain for participating teachers (n = 80).

Table 2
CLASS domains for two group analysis: Beginning and experienced.

CLASS Domain	Teacher Experience Category		$F(1, 77)$	p
	Beginning (0–3 years) $n = 25$ M ± SD	*Experienced* (+3 years) $n = 55$ M ± SD		
Emotional Support	5.33 ± 0.97	5.20 ± .94	.123	.727
Classroom Organisation	4.75 ± 1.36	4.96 ± 1.18	2.063	.155
Instructional Support	3.39 ± 1.19	3.71 ± 1.06	2.489	.119

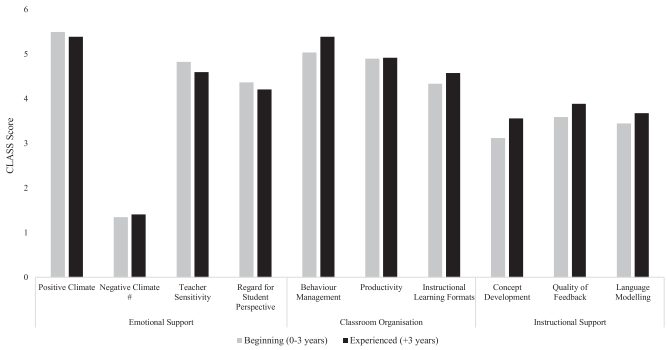

Fig. 2. Comparing mean CLASS scores for Beginning (0–3 years) and Experienced (+3 years) teacher groups for each dimension. There were no statistically significant group differences for any dimension. # Negative Climate is reverse scored.

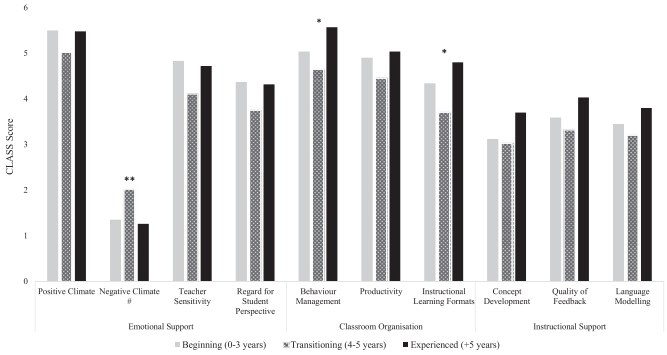

Fig. 3. Comparing mean CLASS scores for Beginning (0–3 years), Transitioning (4–5 years) and Experienced (+5 years) teacher groups for each dimension. # Negative Climate is reverse scored. *p < .05. **p < .01.

3.2. Beginning (0–3 years), transitioning (4–5 years) and experienced teachers (+5 years)

3.2.1. CLASS domains

In the ANCOVA models adjusting for grade level taught, there was a significant difference between the teacher experience categories for the Classroom Organisation domain ($p < .05$; Table 3), but not the Emotional Support and Instructional Support domains. In each domain, regardless of statistical significance, the Beginning and Experienced teacher groups scored higher than the Transitioning group.

3.2.2. CLASS dimensions

The Beginning teacher group had very similar mean scores to the Experienced teacher group (+5 years) in nearly all dimensions of the CLASS, with the Transitioning group (4–5 years) being relatively lower (Fig. 3). Significant group differences were present in Negative Climate ($F(2, 76) = 6.29$, $p = .003$, $\eta_p^2 = 0.14$), Behaviour Management ($F(2, 76) = 4.51$, $p = .014$, $\eta_p^2 = 0.11$), and Instructional Learning Formats ($F(2, 76) = 4.48$, $p = .014$, $\eta_p^2 = 0.11$).

Pairwise comparisons for Negative Climate indicated that the significant difference was in reference to the Transitioning group of teachers. The Transitioning group demonstrated more negative classroom climate behaviours than both the Beginning and Experienced teacher groups ($ps \leq .02$). There was no significant difference between the Beginning and Experienced teacher groups for negative classroom climate.

Pairwise comparisons for Behaviour Management and Instructional Learning Formats dimensions were inconsistent. Pairwise comparisons for the Behaviour Management dimension were not significant. Pairwise comparisons for the Instructional Learning Formats dimension indicated that only the Experienced teacher group had scores that were significantly higher than the Transitioning group ($p < .04$). There were no statistically significant differences with the Beginning teacher group for pairwise comparisons for Instructional Learning Formats.

4. Discussion

This research investigated whether there is any association between teachers' years of experience and teaching quality using CLASS scores from observations of 80 Prep to Grade 3 classroom teachers. Teachers' years of experience were initially categorised into two groups: Beginning (0–3 years) and Experienced (more than 3 years). Our first analysis found no significant differences between Beginning and Experienced teachers on any of the three domains or 10 dimensions measured by the CLASS. These findings are consistent with previous research using the CLASS that found no differences between beginning/novice and experienced teachers (Stuhlman & Pianta, 2009). However, to account for the developmental differences between beginning (or graduate) and early career teachers with more years of experience, we divided our sample into three groups: Beginning (0–3 years), Transitioning

Table 3
CLASS domains for three group analysis: Beginning, transitioning and experienced.

CLASS Domain	Teacher Experience Category			$F(2, 76)$	p	η_p^2
	Beginning (0–3 years) $n = 25$ M ± SD	*Transitioning* (4–5 years) $n = 11$ M ± SD	*Experienced* (+5 years) $n = 44$ M ± SD			
Emotional Support	5.34 ± 0.97	4.72 ± 1.39	5.32 ± .76	1.604	.208	.04
Classroom Organisation	4.76 ± 1.36	4.26 ± 1.40	5.14 ± 1.07	3.296	.042	.08
Instructional Support	3.39 ± 1.19	3.18 ± 1.34	3.84 ± .96	2.60	.081	.06

(4–5 years), and Experienced (more than 5 years). In this second analysis, a significant main effect of group was found for Classroom Organisation. The Classroom Organisation domain difference reflected differences at the dimension level with Beginning and Experienced teachers outperforming teachers in the Transitional group in three dimensions: Negative Climate, Behaviour Management, and Instructional Learning Formats.[3]

Neither level of analysis provides evidence to support the claim that beginning teachers (with 0–3 years' experience) are inadequately prepared for the classroom and may instead provide further evidence of a decline in the quality of teaching after the first three years for some teachers (Chingos & Peterson, 2011; Klassen & Chiu, 2010; Rivkin et al., 2005; Rockoff, 2004). The removal of initial support structures, together with an increase in workload and responsibilities, and emergence of emotional burnout may contribute to this decline (Buchanan et al., 2013; Weldon, 2018). Previous research investigating associations between teaching quality and years of experience has also questioned whether a previously detected post-five-year improvement is associated with early career teacher attrition and the departure of less effective early career teachers (Chingos & Peterson, 2011; Clotfelter, Ladd, & Vigdor, 2006). Taken together, these findings suggest that the assessment of beginning teachers' performance could be negatively impacted by the use of broad experience categories that extend beyond the initial 0–3 year phase to potentially include what several studies, including this one, have indicated may be a transitional period and one that may be difficult for some early career teachers. Further research with a larger number of transitional-stage teachers from a range of different cultural, school phase, and socioeconomic contexts is required before generalising the findings reported here.

In the meantime, however, our findings do suggest that placing emphasis on the accreditation, content, and availability of high-quality mentoring and ongoing professional learning for *all* teachers may be more appropriate than further experiments in ITE pathways and the destabilisation of university teacher education. This observation is based on the mean scores achieved by our full sample which, while not representative of all Australian teachers and including only primary school teachers, is comparable with research conducted using the CLASS internationally. This research points to an established pattern of higher scores in the Emotional Support and Classroom Organisation domains, together with lower scores in the Instructional Support domain (Hamre & Pianta, 2009). The distribution of scores in our sample for each of these domains is also similar to that of Head Start grantees in the United States (Head Start Early Childhood Learning & Knowledge Center, 2016), although there is greater spread in scores in this study indicating more individual variability in teaching quality. At the same time, however, our results differ to those of Head Start grantees in that there were no mean scores in the high range (6–7) in any of the three CLASS domains. Further, mean scores in the Classroom Organisation domain typically fell below the minimum threshold score (5.00) considered necessary for positive social, academic and behavioural development (The National Center on Quality Teaching and Learning, 2011). By contrast, the mean score of teachers in our sample surpassed the minimum threshold of 3.00 for Instructional Support, which again differs from the Head Start grantee pattern where the mean score was 2.83. A possible explanation for this difference is that Head Start uses the CLASS Pre-K, rather than the CLASS K-3, and the intellectual demand of the Australian Curriculum in the primary school grades may contribute to higher scores in the Instructional Support domain.

A similar pattern was evident at the dimension level with no group achieving high range mean scores in any of the 10 CLASS dimensions. Mean scores for all three groups (Beginning, Transitioning, Experienced) fell below the minimum threshold of 5.00 in two dimensions of Emotional Support (Teacher Sensitivity, Regard for Student Perspectives) and one dimension of Classroom Organisation (Instructional Learning Formats). Further, while our Experienced (more than 5 years) group scored above the threshold in Productivity, the Beginning and Transitioning groups each scored below. The Transitioning group also scored more poorly in Negative Climate than the other two groups and was the only group to score below the threshold in Behaviour Management. One possibility that should not be overlooked and which is worthy of further research is whether cultural differences play a role in the lack of high range mean scores achieved by this sample of Australian teachers, compared to patterns established in the United States. Interestingly, the "enthusiasm" of American teachers and the possibility that some CLASS dimensions, particularly those feeding into the Emotional Support domain, may be "culturally bound" has been raised by European researchers (Bihler, Agache, Kohl, Willard, & Leyendecker, 2018, p. 3). The possibility of a similar effect of culture should be examined here.

Despite claims that Australian teachers are poor at managing student behaviour, Behaviour Management was the highest scoring dimension for experienced teachers and the second highest scoring dimension, after Positive Climate, for both Beginning and Transitioning teachers with only the Transitioning group scoring below the threshold of 5. This finding suggests that calls for more professional development in behaviour management may be misdirected and is unlikely, by itself, to improve the overall quality of teaching. We say this because teaching is complex and multidimensional. The CLASS observational measure has been designed to take this complexity into account and, as such, there is some interrelatedness between dimensions. For example, low Productivity can play into Behaviour Management (Graham, 2018), as lack of lesson preparation and the absence of smooth transitions can lead to students sitting idle which can prompt talking, fidgeting, and calling out. Some teachers deal with this situation quickly and effectively to move back into teaching. When observed using the CLASS, these teachers would score well in Behaviour Management but lower in Productivity.

However, some teachers can get distracted by these emergent behaviours and respond reactively or ineffectively (Graham, 2018), thereby fuelling the instructional vacuum created by low productivity. These teachers would score poorly in both Productivity and Behaviour Management and, depending on how they attempt to regain control, possibly in Negative Climate as well. Any future initiatives to improve the quality of teaching should match the complexity and multidimensionality of this problem and seek to address behaviour through the implementation of Positive Behaviour Intervention and Supports (PBIS) in schools at a systemic level as part of a comprehensive Multi-Tier Systems of Support (MTSS) framework that also supports academic and social emotional learning (Bohanon, Gilman, Parker, Amell, & Sortino, 2016; Yeung et al., 2016). This is not something that can be achieved through changes to ITE alone, although mandating inclusion of at least one stand-alone unit within the ITE curriculum to enable coverage of the principles of MTSS and PBIS, as well as evidence-based classroom management practices would be a welcome change that might prevent this important area of practice being embedded or lost to a crowded curriculum.

These recommendations are supported by our results which, in

[3] The Behaviour Management and Instructional Learning Format dimensions measure teachers' proactive use of different strategies, modalities and materials to engage students and to keep them on task. Negative Climate increases when teachers' use punitive methods to achieve control, such as threats or sarcasm.

finding no additional return to experience and rather some evidence of a decline after three years, align with the bulk of previous empirical research investigating teachers' years of experience and the quality of teaching. To date, however, this empirical evidence appears not to have been consulted in the construction of policies attempting to "vet" potential entrants to the teaching profession through changes to and greater prescription of university ITE programs. Together, these policies and their accompanying rhetoric have positioned beginning teachers as "the problem" affecting student engagement, achievement and behaviour, and reform of university teacher education as "the solution". However, as there are now multiple studies using both direct and indirect measures that have found no significant difference in the quality of teaching between beginning teachers and teachers with many more years of experience, continuation of this framing suggests that political calculation takes higher priority in the formulation of education policy and workforce development than empirical evidence. While attracting, educating and retaining high quality candidates into the teaching profession is of critical importance, our data suggest that focusing purely on beginning teachers and their preparation is not the only "solution" to teaching quality, and that any future initiatives should aim to improve the capacity for quality teaching of *all* teachers, especially those who may be experiencing emotional burnout.

4.1. Limitations

A potential limitation in this analysis is the number of teachers ($n = 11$) in the Transitioning group (4–5 years' experience). With the exception of Negative Climate, which is reverse scored and was significantly higher for the Transitioning group, the Transitioning group typically had lower and more variable CLASS scores than the Beginning and Experienced teachers. Further, as presented in Table 3, the Transitioning group means typically had the largest standard deviations, which may indicate individual outliers. This warrants further investigation with a larger number of teachers with 4–5 years' experience.

Another potential limitation is that the CLASS measure was developed and normed in the United States. The possibility that some CLASS dimensions may be "culturally bound" has been raised in research conducted in Europe where, for example, German teachers scored lower as a group in some dimensions, than American teachers. As a group, Australian teachers may be more emotionally reserved than their American counterparts, however, many individual teachers in this sample also did achieve scores in the high range suggesting that high levels of positivity are possible.

Finally, this study included only primary school teachers using the CLASS K-3 measure. Recognising that developmental and curriculum differences have an influence on classroom pedagogy, two other versions of the CLASS were developed for the Upper Elementary (Grades 5 and 6), and Secondary (Grades 7–12) school phases. Further research is necessary to assess whether there is a relationship between teachers' years of experience and the quality of teaching in the middle and secondary school years.

5. Conclusion

This study used standardised classroom observations from 80 Queensland primary school teachers teaching Prep to Grade 3 to examine associations between teachers' years of experience and the quality of teaching. Teaching quality was assessed using the Classroom Assessment Scoring System (CLASS) and results were compared across two, and then three, experience categories based on the extant literature. Our analyses found no evidence that Beginning teachers (0–3 years' experience) were less competent than teachers with more years of experience at either the domain or dimension level measured by the CLASS. However, differences were detected for teachers in the Transitioning group (4–5 years' experience) who achieved significantly lower scores in the Classroom Organisation domain. At the dimension level, the Transitioning group also scored significantly lower in Negative Climate, Behaviour Management, and Instructional Learning Format.

Contrary to reports suggesting Australian teachers are poor at managing behaviour, Behaviour Management was among the highest scoring dimensions for all three groups, behind only Positive Climate for Beginning and Transitioning teachers. The distribution of scores for teachers in the full sample was not unlike those reported in the international research literature, with higher scores in Emotional Support and Classroom Organisation, than Instructional Support. Unlike some international samples, however, no groups achieved mean scores in the high range for any domain or dimension of the CLASS. While teachers in this sample typically achieved scores above the minimum threshold considered necessary for students to make positive gains in the social, emotional, and academic domains of learning, the results were not outstanding.

The findings from this research suggest that Beginning teachers are doing as well or better than teachers with more years of experience, but that the overall quality of teaching could be higher. As such, better support and professional learning is necessary to improve the quality of teaching for *all* teachers, not just those at the beginning of their career. Rather than participate in "formal, one-size-fits-all 'in-service' or conference-style" (Mockler, 2020, p. 4) professional development, it may be that beyond the first years of teaching (0–3 years), and particularly during the 4–5 year transition period, that professional support needs to become more individualised. For example, identifying those teachers with low to mid-range CLASS scores ($\leq$4) for targeted professional support that is relevant to the individual and to the grade level they teach. Strengthening this transition may result in greater teacher satisfaction and retention, and higher levels of teaching quality as these teachers develop greater expertise and confidence in their teaching.

To help produce a more responsive, targeted and effective professional learning and support culture for *all* teachers, not just those at the beginning of their careers, members of the teaching profession should explain to those policy actors with the power to make decisions that: (i) teaching is a complex and multidimensional activity underpinned by ongoing professional learning and support, (ii) more years of experience do not by themselves translate into higher quality practice, (iii) beginning teachers can and do engage in quality teaching at the level of more experienced teachers, and that (iv) some beginning teachers have difficulty transitioning past the initial induction/support period, and that this needs to be identified and addressed sensitively and comprehensively with individually targeted support.

Author contribution

Linda J. Graham: Conceptualisation, Funding acquisition, Project administration, data collection, analysis, writing – review and original draft preparation. Sonia L. J. White: Data curation, quantitative data analysis, writing – review and original draft preparation. Kathy Cologon: Writing – review and original draft preparation. Robert C. Pianta: Analytical strategy, writing and original draft preparation.

Acknowledgements

The first two years of this six-year longitudinal study were

funded by the Australian Financial Markets Foundation for Children (2013-030). The final four years of the study were supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (DP160100319). The views expressed herein are those of the authors and are not necessarily those of the Australian Government or the Australian Research Council.

References

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *Quarterly Journal of Economics, 131*(3), 1415–1453. https://doi.org/10.1920/wp.ifs.2016.0916

Baxter, J. A. (2013). Professional inspector or inspecting professional? Teachers as inspectors in a new regulatory regime for education in England. *Cambridge Journal of Education, 43*(4), 467–485. https://doi.org/10.1080/0305764X.2013.819069

Bihler, L. M., Agache, A., Kohl, K., Willard, J. A., & Leyendecker, B. (2018). Factor analysis of the classroom assessment scoring system replicates the three-domain structure and reveals no support for the bifactor model in German preschools. *Frontiers in Psychology, 9*, 1232. https://doi.org/10.3389/fpsyg.2018.01232

Bohanon, H., Gilman, C., Parker, B., Amell, C., & Sortino, G. (2016). Using school improvement and implementation science to integrate multi-tiered systems of support in secondary schools. *Australasian Journal of Special Education, 40*, 99–116. https://doi.org/10.1017/jse.2016.8

Bolton, R. (2019, July 25). Ideology crushes teachers' ability to control classes. *Australian Financial Review*. Retrieved from https://www.afr.com/policy/health-and-education/ideology-crushes-teachers-ability-to-control-classes-20190724-p52a6s.

Brandenburg, R., McDonough, S., Burke, J., & White, S. (Eds.). (2016). *Teacher education: Innovation, intervention and impact*. Springer. https://doi.org/10.1007/978-981-10-0785-9.

Buchanan, J., Prescott, A., Schuck, S., Aubusson, P., Burke, P., & Louviere, J. (2013). Teacher retention and attrition: Views of early career teachers. *Australian Journal of Teacher Education, 38*, 112–129. https://doi.org/10.14221/ajte.2013v38n3.9

Carey, A. (2019, December 4). *Australian students 'among the worst in the world' for class discipline. The Sydney Morning Herald*. Retrieved from https://www.smh.com.au/politics/federal/australian-students-among-the-worst-in-the-world-for-class-discipline-20191204-p53gwk.html.

Chingos, M. M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review, 30*(3), 449–465. https://doi.org/10.1016/j.econedurev.2010.12.010

Churchward, P., & Willis, J. (2019). The pursuit of teacher quality: Identifying some of the multiple discourses of quality that impact the work of teacher educators. *Asia-Pacific Journal of Teacher Education, 47*(3), 251–264. https://doi.org/10.1080/1359866X.2018.1555792

Clotfelter, T., Ladd, H. F., & Vigdor, J. L. (2006). *Teacher-student matching and the assessment of teacher effectiveness*. Cambridge, MA: National Bureau of Economic Research Working Paper No. 11936. https://doi.org/10.3386/w11936

Cortina, K. S., Miller, K. F., McKenzie, R., & Epstein, A. (2015). Where low and high inference data converge: Validation of CLASS assessment of mathematics instruction using mobile eye tracking with expert and novice teachers. *International Journal of Science and Mathematics Education, 13*(2), 389–403. https://doi.org/10.1007/s10763-014-9610-5

Flores, M. A. (2019). Unpacking teacher quality: Key issues for early career teachers. In A. Sullivan, B. Johnson, & M. Simons (Eds.), *Attracting and keeping the best teachers* (pp. 15–38). Singapore: Springer.

Gale, T., & Parker, S. (2017). The prevailing logic of teacher education: Privileging the practical in Australia, England and Scotland. In M. A. Peters, B. Cowie, & I. Menter (Eds.), *A companion to research in teacher education* (pp. 521–535). Singapore: Springer. https://doi.org/10.1007/978-981-10-4075-7_35.

Gallant, A., & Riley, P. (2014). Early career teacher attrition: New thoughts on an intractable problem. *Teacher Development, 18*(4), 562–580. https://doi.org/10.1080/13664530.2014.945129

Gore, J. M. (2016). Reform and the reconceptualisation of teacher education in Australia. In R. Brandenburg, S. McDonough, J. Burke, & S. White (Eds.), *Teacher education* (pp. 15–34). Singapore: Springer. https://doi.org/10.1007/978-981-10-0785-9_2.

Graham, L. J. (2018). Student compliance will not mean "all teachers can teach": a critical analysis of the rationale for "no excuses" discipline. *International Journal of Inclusive Education, 22*(11), 1242–1256. https://doi.org/10.1080/13603116.2017.1420254

Hamre, B. K., Hatfield, E, Pianta, R, & Jamil, F (2014). Evidence for general and domain-specific elements of teacher–child interactions: Associations with preschool children's development. *Child Development, 85*(3), 1257–1274. https://doi.org/10.1111/cdev.12184

Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). *Building a science of classrooms: Application of the CLASS framework in over 4,000 US early childhood and elementary classrooms*. New York: Foundation for Child Development.

Head Start Early Childhood Learning & Knowledge Center. (2016). *A national overview of grantee CLASS® scores in 2016*. Retrieved from https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/national-class-2016-data.pdf.

Ingvarson, L., & Rowe, K. (2008). Conceptualising and evaluating teacher quality: Substantive and methodological issues. *Australian Journal of Education, 52*(1), 5–35. https://doi.org/10.1177/000494410805200102

Jerrim, J., & Vignoles, A. (2016). The link between East Asian 'mastery' teaching methods and English children's mathematics skills. *Economics of Education Review, 50*, 29–44. https://doi.org/10.1016/j.econedurev.2015.11.003

Joseph, B. (2017). *Getting the most out of Gonski 2.0: The evidence base for school investments*. Sydney, NSW: Centre for Independent Studies. Retrieved from https://www.cis.org.au/app/uploads/2017/10/rr31-e.pdf.

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. https://doi.org/10.1037/a0019237

La Paro, K. M., Pianta, R. C., Stuhlman, M., & Hamre, B. (2002). *Classroom assessment scoring system (CLASS)*. Charlottesville, VA: National Center for Early Development and Learning. University of Virginia.

Ladd, H. F., & Sorensen, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy, 12*(2), 241–279. https://doi.org/10.1162/EDFP_a_00194

Lauen, D. L., & Gaddis, S. M. (2016). Accountability pressure, academic standards, and educational triage. *Educational Evaluation and Policy Analysis, 38*(1), 127–147. https://doi.org/10.3102/0162373715598577

Lingard, B., & Sellar, S. (2019). Schooling reform in Australia: Legitimation through "projections" onto Shanghai and east asian schooling systems. In F. Waldow, & G. Steiner-Khamsi (Eds.), *Understanding PISA's attractiveness: Critical analyses in comparative policy studies* (pp. 49–64). London, UK: Bloomsbury Publishing.

Mayer, D., Dixon, M., Kline, J., Kostogriz, A., Moss, J., Rowan, L., & White, S. (2017). Studying the effectiveness of teacher education. In D. Mayer, M. Dixon, J. Kline, A. Kostogriz, J. Moss, & L. Rowan (Eds.), *Studying the effectiveness of teacher education* (pp. 13–26). Singapore: Springer. https://doi.org/10.1007/978-981-10-3929-4_2.

McGraw, A., & Fish, T. (2018). Selection and rejection in teacher education: Qualities of character crucial in selecting and developing teacher education students. *Asia-Pacific Journal of Teacher Education, 46*(2), 120–132. https://doi.org/10.1080/1359866x.2017.1355048

McMahon, M., Forde, C., & Dickson, B. (2015). Reshaping teacher education through the professional continuum. *Educational Review, 67*(2), 158–178. https://doi.org/10.1080/00131911.2013.846298

Mockler, N. (2018). Early career teachers in Australia: A critical policy historiography. *Journal of Education Policy, 33*(2), 262–278. https://doi.org/10.1080/02680939.2017.1332785

Mockler, N. (2020). *Teacher professional learning under audit: Reconfiguring practice in an age of standards*. Professional Development in Education. https://doi.org/10.1080/19415257.2020.1720779

Morgan, A., & Aspland, T. (2018, September 8). *Viewpoints: Should universities raise the ATAR required for entrance into teaching degrees? The conversation*. Retrieved from https://theconversation.com/viewpoints-should-universities-raise-the-atar-required-for-entrance-into-teaching-degrees-102841.

Nsw Government Department of Customer Service. (2019, September 3). *Ensuring excellence new standards for graduate teachers*. Retrieved from https://www.nsw.gov.au/your-government/the-premier/media-releases-from-the-premier/ensuring-excellence-new-standards-for-graduate-teachers/.

Organisation for Economic Co-operation and Development. (2019). *September 10). Review education policies: Learning time and disciplinary climate*. Retrieved from https://gpseducation.oecd.org/revieweducationpolicies/#!node=41718&filter=all.

O'Neill, S., & Stephenson, J. (2012). Does classroom management coursework influence pre-service teachers' perceived preparedness or confidence? *Teaching and Teacher Education, 28*(8), 1131–1143. https://doi.org/10.1016/j.tate.2012.06.008

Palmer, D. J., Stough, L. M., Burdenski, T. K., Jr., & Gonzales, M. (2005). Identifying teacher expertise: An examination of researchers' decision making. *Educational Psychologist, 40*(1), 13–25.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109–119. https://doi.org/10.3102/0013189X09332374

Podolsky, A., Kini, T., & Darling-Hammond, L. (2019). Does teaching experience increase teacher effectiveness? A review of US research. *Journal of Professional Capital and Community, 4*(4), 286–308. https://doi.org/10.1108/JPCC-12-2018-0032

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*(3), 700. https://doi.org/10.1037/a0027268

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247–252. https://doi.org/10.1257/0002828041302244

Rucinski, C. L., Brown, J. L., & Downer, J. T. (2018). Teacher–child relationships, classroom climate, and children's social-emotional and academic development. *Journal of Educational Psychology, 110*(7), 992. https://doi.org/10.1037/edu0000240

Sachs, J. (2016). Accountability, standards and activism: A challenge or opportunity for teacher education. In C. Day, & J. C. Lee (Eds.), *Quality and change in teacher education: Western and Chinese perspectives* (pp. 251–262). Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-24139-5_15.

Schachter, R. E., Spear, C. F., Piasta, S. B., Justice, L. M., & Logan, J. A. (2016). Early childhood educators' knowledge, beliefs, education, experiences, and children's language- and literacy-learning opportunities: What is the connection? *Early Childhood Research Quarterly, 36*, 281–294. https://doi.org/10.1016/j.ecresq.2016.01.008

Scholes, L., Lampert, J., Burnett, B., Comber, B. M., Hoff, L., & Ferguson, A. (2017). The politics of quality teacher discourses: Implications for pre-service teachers in high poverty schools. *Australian Journal of Teacher Education, 42*(4), 19–43. https://doi.org/10.14221/ajte.2017v42n4.3

Sellar, S., Thompson, G., & Rutkowski, D. (2017). *The global education race: Taking the measure of PISA and international testing*. Edmonton, Alberta: Brush Education.

Stuhlman, M. W., & Pianta, R. C. (2009). Profiles of educational quality in first grade. *The Elementary School Journal, 109*(4), 323–342. https://doi.org/10.1086/593936

Sullivan, A., Johnson, B., & Simons, M. (Eds.). (2019). *Attracting and keeping the best teachers: Issues and opportunities*. Singapore: Springer.

Teacher Education Ministerial Advisory Group. (2014). *Action now: Classroom ready teachers*. Canberra, Australia: Australian Government, Department of Education. Retrieved from https://docs.education.gov.au/system/files/doc/other/action_now_classroom_ready_teachers_print.pdf.

The National Center on Quality Teaching and Learning. (2011). *Understanding and using the CLASS for program improvement*. Retrieved from http://prekkid.org/head-start/agency/training-and-tech-assistance/Class%20Overview%20Materials/NCQL%20Understanding%20and%20Using%20the%20CLASS%20%5BHappy%20Tsugawa-Banta%5D.pdf.

Voss, T., Wagner, W., Klusmann, U., Trautwein, U., & Kunter, M. (2017). Changes in beginning teachers' classroom management knowledge and emotional exhaustion during the induction phase. *Contemporary Educational Psychology, 51*, 170–184. https://doi.org/10.1016/j.cedpsych.2017.08.002

Weldon, P. (2018). Early career teacher attrition in Australia: Evidence, definition, classification and measurement. *Australian Journal of Education, 62*(1), 61–78. https://doi.org/10.1177/0004944117752478

Yeung, A. S., Craven, R. G., Mooney, M., Tracey, D., Barker, K., Power, A., & Lewis, T. J. (2016). Positive behavior interventions: The issue of sustainability of positive effects. *Educational Psychology Review, 28*(1), 145–170. https://doi.org/10.1007/s10648-015-9305-7