Received: 24 June 2022 | Accepted: 7 November 2023

DOI: 10.1002/berj.3938

ORIGINAL ARTICLE

BERJ BERA

Can teacher quality be profiled? A cluster analysis of teachers' beliefs, practices and students' perceptions of effectiveness

Michael Witter[1] | **John Hattie**[2]

[1]Faculty of Education and Arts, Australian Catholic University, Melbourne, Australia

[2]Melbourne Graduate School of Education, Melbourne, Australia

Correspondence
Michael Witter, Faculty of Education, School of Education and Arts, Australian Catholic University, Office 460.1.51 Level 1, 250 Victoria Pd, East Melbourne VIC 3002, Locked Bag 4115, Fitzroy, VIC 3065, Australia.
Email: michael.witter@acu.edu.au

Present address
Michael Witter, Saint Teresa of Kolkata Building, 421 - 115B Victoria Parade, Level 9 - Room 9.14, Fitzroy, VIC 3065, Australia

Abstract

A critical aim of teacher quality research is to identify and distil its multiple components. While many such components have been identified in the literature, there is a scarcity of research that has attempted to empirically investigate how these elements integrate within a larger profile of teacher quality, including how particular teachers' beliefs, behaviours and outcomes cluster into unique profiles of teachers. In this study, a group of Australian high-school teachers completed an extensive teacher beliefs questionnaire, undertook the Visible Classroom program to record and document their use of high-leverage teaching practices and had their students complete a survey based on perceptions of effective teaching. The results were analysed using cluster analysis to determine if distinct groups of teachers could be identified based on similarities in beliefs, behaviours and student perceptions of teaching. The results suggested multiple distinct clusters of teachers with distinguishing types of beliefs, frequency of practices and student perceptions of teaching. Most notably, the results pointed to a unique cluster of teachers who were most distinguishable in their self-efficacy, intrinsic motivation and personal responsibility beliefs, as well as highly active teaching behaviours in the classroom. This group also possessed the highest levels of student perceptions of teaching. The findings provide new insights into the complex profiles that shape teacher quality and implications for future research.

KEYWORDS

approaches to teaching, evaluation, survey, teacher research

Key insights

What is the main issue that the paper addresses?

The paper investigates whether a multidimensional profile of teacher quality can be identified through cluster analysis, with particular emphasis on teachers' beliefs, their use of empirically supported high-leverage teaching practices and their students' perceptions of teaching.

What are the main insights that the paper provides?

Distinct clusters of teachers can be identified based on particular beliefs, frequency of high-leverage teaching practices and students' perceptions of teaching. A profile of teacher quality emphasised higher levels of self-efficacy, intrinsic motivation and personal responsibility, coupled with active teaching and high levels of student perceptions of quality teaching.

INTRODUCTION

There is general consensus among researchers that teachers play an essential role in shaping learner outcomes. It is thus unsurprising that teacher quality has consistently been and remains a focus of research and policymaking across education systems worldwide.

A critical question at the heart of this global prioritisation of teaching is *what makes a great teacher?* Answers to this question have evolved significantly over the decades, as the knowledge base and research methods have grown in efforts to identify, measure and describe the characteristics, behaviours and outcomes of effective teachers. While definitions and models of teacher quality are not consistently agreed upon, few researchers would disagree that teacher quality is complex and multifaceted, which is in part a reflection of changing societal views and demands on the role of the teacher and expectations for quality within this role (Muijs, 2006).

Addressing the question of what makes a great teacher then has become an exercise in identifying multiple critical elements and situating these within a larger, integrated whole: a profile of teacher quality. Policies and practices dedicated to improving teaching quality hinge on implicit conceptualisations of what it means to be a quality teacher. Developing profiles of teacher quality that privilege its multidimensional nature and seek to examine the integration of its various elements is therefore essential. Yet, how we do this remains uncertain and underexamined by the extant research. Teacher quality and teaching quality, although related and often conflated, are distinct (Kennedy, 2010). Teaching quality, which concerns the practices of teachers, forms just one of several elements of teacher quality. Goe's (2007) framework for teacher quality demarcates three key elements: inputs, processes and outcomes. Inputs comprise characteristics such as prior education, experience, attributes, beliefs and gender. Processes include classroom practices such as classroom management and other activities such as planning. Outcomes relate to student achievement gains based on teachers' value-added. If teacher quality is reflected through a range of integrated inputs, processes and outcomes, then presumably one should be able to create a profile (or profiles) of teacher quality based on multiple elements. Although there are examples within the literature that have attempted to put forward such profiles (e.g., see Breault, 2013; Colker, 2008; Stronge, 2018), these are often done via a composite approach

to reviewing potential teacher quality factors (see Darling-Hammond, 2000; Goe, 2007; Rice, 2003; Wayne & Youngs, 2003). Actual empirical studies attempting to construct and examine multifaceted profiles of multiple integrated teacher quality elements are relatively scarce.

This study sought to empirically investigate whether a profile or profiles of teacher quality could be distilled based on a range of factors, focusing on five higher-order factors comprised of 31 distinct and validated types of teacher belief, two higher-order types of effective teaching practice based on 14 effective teaching practices empirically supported by meta-analytic research and two aspects of instructional effectiveness identifiable via student perception surveys. The ultimate aim was to determine whether particular clusters of teachers would unify around a set of common beliefs and behaviours, aligned with high perceptions of teaching quality on the part of their students. If this were the case, it could indicate a particular profile encompassing multiple dimensions of teacher quality. The results of the study suggested that there may be multiple distinct profiles of teachers with relatively common beliefs and behaviours, and that students may perceive teachers with distinctly different (and in some ways contradictory) profiles as effective in the classroom. Yet, among the different clusters, there was one particular group that stood out in terms of their combined beliefs, behaviours and their students' perceptions of teaching (which were among the highest in the study), a group that offered a possible profile of teacher quality.

ESTABLISHING A PROFILE OF TEACHER QUALITY

It is difficult to establish a profile of teacher quality as one views the complexities of the profession itself. On the one hand, Hanushek (2002) argued quite simply that 'Good teachers are ones who get large gains in student achievement for their classes; bad teachers are just the opposite' (p. 4). On the other hand, critics of this view argue that such views force narrow and inaccurate conceptualisations of teacher quality, and disregard the realities and complexities of the schools, classrooms and professional requirements of teachers (Cochran-Smith, 2003). Cochran-Smith urged educators to embrace these complexities and pursue accountability for teacher quality via more intricate and multifaceted conceptualizations of the construct.

In the spirit of the latter view of teacher quality, research profiles of teacher quality are most often constructed via literature reviews of unique factors, which are effectively listed or drawn together. There is little common consensus among researchers when it comes to determining what criteria and methods ought to be used to establish such a profile via such methods. Illustrative examples of this can be found in the work of Brophy and Good (1984), Brophy (1986), Goe (2007) and Stronge (2018). Brophy and Good (1984) reviewed over a decade of research linking teacher behaviours to student outcomes, focusing primarily on process–product research linking presumably causal factors to student achievement gains. This resulted in a profile primarily of teacher practices they associated with quality teaching, noting, in particular, the importance of 'active teaching' in effective classrooms (where students spend most of their time being taught or closely supervised by teachers), high-quality classroom management, brisk pacing, teacher clarity when delivering instruction and the teacher's ability to create high success rates for learners within lessons. The authors also noted some qualities that expanded beyond teacher actions, including the value of higher expectations for students, an orientation towards focusing on instruction and learning activities and a business-like approach to teaching that emphasised time on task.

Stronge (2018) presented a profile centred on ten qualities of effective teachers, which the authors described as a synthesis of research drawn from a range of studies, including previous literature reviews, meta-analyses, product–process research, case studies, surveys,

interviews and other forms of qualitative research. According to the authors, among other things, effective teachers are formally trained and certified within their fields, experienced, organised, respectful and caring, effective planners, strong classroom managers and able to present content in engaging ways, monitor learning, provide feedback and address the needs of diverse learners.

Another profile, from Goe et al. (2008), aimed to update studies conducted in prior years, such as those by Darling-Hammond (2000), Rice (2003) and Wayne and Youngs (2003). The authors constructed a five-point profile of teacher quality based on recent and seminal reviews of the literature, tested and then revised it with input from experts on teacher quality and effectiveness. They suggested that effective teachers have high expectations, produce improved student achievement gains, as well as contribute to social/emotional learning outcomes, plan and structure engaging learning opportunities, monitor student progress and adapt as needed, promote diversity and inclusion and collaborate effectively with other teachers, staff and families.

These examples reflect several general features relevant to researchers' attempts to build teacher quality profiles. The first key feature is that such profiles do not comprehensively represent the range of practices empirically supported as effective at raising student achievement gains, but rather attempt to distil a limited number of general, presumably high-leverage, teaching practices. The practices emphasised by different authors overlap, but also vary, and are borne out of a multitude of studies providing a solid empirical basis in support of these practices. Brophy and Good's research was a seminal synthesis of process–product research, while Stronge and Tucker's profile is distilled from dozens of studies, including process–product research, literature reviews and qualitative studies. Goe et al. distilled a profile of teacher quality by building upon recent existing reviews, coupled with an iterative process involving input from experts in the field.

Although the profiles illustrated by these key publications are well supported by the literature, they also reflect a persistent challenge in identifying the practices of quality teachers. If claims around positive influences on learner outcomes are based on their positive correlation to improved student achievement gains, as is generally the case, then nearly everything teachers do can be claimed as effective practice. Hattie (2009) found that between 90% and 95% of all influences on student achievement are positive. In essence, as Hattie put it: 'Virtually everything works. One only needs a pulse' (p. 16). Profiles that identify a discrete and non-comprehensive set of teaching practices essentially privilege certain practices over others, including others that may be effective or valued based on a particular set of criteria, including but not limited to the degree of impact on student learning associated with that practice. Consequently, most models of quality teaching can be viewed simultaneously as valid and trivial, depending on the criteria used to define and identify key practices. The challenge Hattie raised is not in determining what works in terms of such practices, but what works best. While Hattie argued that average effect sizes of teacher influences can be utilised to differentiate between low, medium and high influences, with a hinge point of $d = 0.4$ to indicate where influences (including practices) fall into the zone of desired effects, this approach has not been widely applied to or accepted within teacher quality research as a way to determine not just which practices may set apart quality teachers from their peers, but also which practices are most central to their effectiveness.

DESIGNING RESEARCH THAT PROFILES TEACHER QUALITY

It may be the case that utilising reviews of teacher quality characteristics to synthesise a profile of teacher quality requires a logical leap beyond the actual scope of the studies

upon which those reviews are based. It is rarely if ever the case that the full scope of characteristics addressed in these reviews has been examined within single studies, and thus it is problematic to assume that one can simply cobble together a profile of a quality teacher via literature reviews, disregarding how those characteristics may integrate in actual practice. For this reason, it would be important to look at studies that have attempted to integrate a range of characteristics in the interests of establishing a teacher quality profile. The majority of studies that focus on a profile of integrated characteristics tend to emphasise constructivist approaches with limited generalisability, or focus on measures that are not direct indicators of quality teaching. For example, Gay (2012) identified six key dimensions of quality mathematics teachers using a collective case study approach with three teachers, while Colker (2008) interviewed 43 early childhood practitioners about their perceptions of the characteristics of effective educators, analysing the results and identifying 12 characteristics that trended among participants. Keeley et al. (2016) had a group of award-winning tertiary educators rate each of the teaching behaviours on a checklist based on importance, using these to determine the most important characteristics. Such studies provide valuable insights on quality teaching from the perspective of teachers, but lack the kind of validation methodologies that enable researchers to test some of the findings regarding multiple facets of teacher quality provided in reviews of the literature within single studies.

There are a few notable studies that have used empirical methodologies to provide a kind of profile of teacher quality that draws upon a range of integrated factors. For instance, Guo et al. (2012) used structural equation modelling to examine a range of factors, including teacher experience, teacher education, teacher self-efficacy beliefs, teaching practices and student learning outcomes in English, to determine which factors (if any) contributed directly or indirectly to learner outcomes. The results produced an implicit profile of teacher quality that notably included higher levels of self-efficacy and more frequent positive interactions between teachers and their students, positive feedback, productive use of instructional time and responsiveness to students during instruction. The profile ultimately failed to provide support for other factors such as teacher education, experience and time spent on content. Stronge et al. (2011) used cross-case analysis to first distinguish groups of effective teachers based on value-added to student achievement, and then examine a range of factors, including teacher self-efficacy, teaching practices such as questioning techniques, their students' time on task and an observational rating form. This approach enabled the authors to build a profile of quality teachers that emphasised their depth of expertise in a range of classroom management techniques, a higher degree of fairness and respect, as well as more positive relationships with students, but they also could not find significant differences between effective and ineffective teachers based on their instructional delivery and assessment practices, nor were these teachers distinguishable based on their self-efficacy beliefs.

Stronge et al. cautioned that the results of studies like these should not be treated as silver-bullet indicators of teacher effectiveness, emphasising that 'effective teaching involves a dynamic interplay among content, pedagogical methods, characteristics of learners, and the contexts in which the learning will occur' (p. 349). This perspective illustrates the importance of further studies of this nature. A major gap in the current research is the over-reliance on literature reviews in developing teacher quality profiles, signalling a critical need for a larger body of empirical research that examines multiple integrated aspects of teacher quality, putting the literature to the test. It was in light of this need for further research that this study was conducted.

METHOD

Participants

Participants were current or former participants of the Teach For Australia program (www.teachforaustralia.org), teaching in government high schools across the states of Victoria, Australia Capital Territory, Western Australia and the Northern Territory, across multiple subjects and year levels, including English, Social Studies, Science, Mathematics and additional subject areas. 64 of the 68 teachers who volunteered for the study had 1–5 years of experience, and four teachers had 6+ years of experience. 58 of the original 68 volunteers were included in the final analysis, having provided data across the study's three key measures. As high-school teachers, nearly all participants taught multiple distinct subjects, but 35 of 68 most closely identified as English, Arts and/or Social Studies teachers, 26 of 68 most closely identified as Science, Technology and/or Mathematics teachers, and 7 of 69 most closely identified as teachers of languages other than English. There were 1056 student participants as well, who completed anonymous surveys of their teachers as one of the primary measures of this study.

Measures

Teacher beliefs

The first dimension measured was higher-order teacher beliefs factors, via a survey instrument with 139 items organised into 31 beliefs factors within five higher-order categories of teacher belief, validated by Witter (2023a). For each item on the survey, respondents were asked to rate their level of agreement on a six-point scale from 1 (*strongly disagree*) to 6 (*strongly agree*). The instrument was based primarily on pre-validated teacher beliefs survey constructs and related items, except for the generalised teacher expectations scale created by the author. Table 1 provides an overview of the higher-order beliefs examined within the study, and particular types of teacher beliefs integrated within each.

Validity statistics for the original design of this survey can be located in the OSF repository (https://osf.io/mqz3a), which includes factor loadings and reliability statistics for each of the identified factors and scales used in the present study. Table 2 provides the descriptive statistics for the teacher beliefs survey, while Table 3 illustrates the factor structure, loadings and reliability statistics.

Estimates of reliability were sufficiently high to have confidence in the total scores within each of the five identified factors (social/deep $\alpha = 0.86$, responsibility $\alpha = 0.73$, efficacy and intrinsic motivation $\alpha = 0.72$, pro-assessment $\alpha = 0.73$). Caution is needed for surface/fixed as this estimate of reliability ($\alpha = 0.57$) is below acceptable levels. However, the larger sample size and estimates of reliability for the instrument validated by Witter (2023a) are sufficiently high across all five factors, providing additional empirical support for the factor structure fit to the two separate studies.

Student perceptions of teaching

The third dimension of teacher quality, student perceptions of teaching, was measured through a student perception survey. The 35-item Tripod Student Perception survey was devised by Ferguson (Ferguson & Danielson, 2015; LaFee, 2014) and widely utilised in the Measures of Effective Teaching Project (Kane & Staiger, 2012). Each participant

TABLE 1 Higher-order beliefs structure and overview.

Higher-order belief	Descriptive overview	Incorporated teacher belief types
Social/deep	Integrated set of beliefs focused on education for the purposes of improving society, supporting the self-actualisation and individual efficacy of students via teaching, learning and curriculum that emphasises deep knowledge and relational connectedness.	Social reconstructivist curriculum and social reform teaching Prioritising cross-curriculum and individualised teaching and learning Choosing to teach for social utility purposes Deep and relational conceptions of learning Mastery and relational goal orientations
Surface/fixed	Set of beliefs that broadly combine fixed mindset characteristics about self and students, coupled with surface and transmission-driven conceptions of teaching, curriculum and learning.	Belief that student ability is largely innate Rejection of the importance of effort versus ability Privileging knowledge that is simple Choosing to teach for personal utility value and professional status Work-avoidant and ability-focused goal orientation Focus on curriculum and teaching in systematic and transmissive ways Low expectancy beliefs about students
Responsibility	Set of beliefs reflecting four dimensions of teacher personal responsibility.	Personal responsibility for the achievement of students Personal responsibility for the motivation of students Personal responsibility for the quality of teaching Personal responsibility for relationships with students
Pro-assessment	Set of beliefs that reflect a positive orientation towards assessment.	Belief that assessment creates positive accountability for students and schools Belief that assessment helps demonstrate learning Belief that assessment supports improved teaching and learning Belief that assessment does not hinder teaching and learning
Efficacy and intrinsic motivation	Set of self-beliefs in the ability to engage in a variety of important instructional practices in the classroom and motivation to teach based on these self-beliefs and a general intrinsic value of teaching.	Self-efficacy of classroom management Self-efficacy of instruction Choosing to teach based on intrinsic motivation and self-efficacy Pursuing and self-reporting teaching for deep learning
Learning	Learning—deep.	Belief that learning is building deep knowledge that can be applied to the real world

TABLE 2 Teacher beliefs: descriptive statistics.

Belief type	*N*	Mean	SD
Learning—deep (widening understanding and real-world application)	68	5.52	0.50
Purpose—social utility value	67	5.47	0.60
Mastery—goal orientation approach	68	5.37	0.61
General expectations for achievement	67	5.26	0.49
Curriculum—cross-curricular skills	67	5.21	0.71
Responsibility—relationships	67	5.13	0.82
Teaching for deep learning	68	5.03	0.66
Learning—deep (relational)	68	5.02	0.74
Relational goal orientation	68	4.97	0.83
Assessment—improvement	68	4.94	0.87
Purpose—intrinsic motivation and self-efficacy	67	4.81	0.83
Responsibility—teaching	67	4.74	0.96
Curriculum—social reconstruction	67	4.73	0.90
Social reform teaching	68	4.71	0.87
Curriculum—individualisation	67	4.59	0.93
Self-efficacy for classroom management	68	4.44	0.80
Self-efficacy for instructional practice	68	4.28	0.69
Responsibility—student achievement	67	4.14	1.00
Responsibility—student motivation	67	4.04	1.07
Assessment—demonstrates learning	68	3.90	0.86
Curriculum—systematic	67	3.69	0.79
Assessment—accountability	68	3.36	0.77
Ability—goal orientation approach	68	3.29	1.01
Assessment—hindrance	68	3.12	0.81
Learning—surface	68	2.87	0.89
Purpose—personal utility value and professional status	67	2.18	0.81
Work—goal orientation avoidance	68	2.00	0.71
Innate ability	68	1.91	0.67
Ability—goal orientation avoidance	68	1.69	0.56
Epistemology—simple	68	1.52	0.47
Importance of student effort	68	1.22	0.31
	N	**Mean**	**SD**
Learn_Deep	68	5.52	0.50
Purpose_SUV	67	5.47	0.60
GO_MasteryAppr	68	5.37	0.61
Expectations	67	5.26	0.49
Curriculum_CrossCurric	67	5.21	0.71
Responsibility_Relationships	67	5.13	0.82
Teach_Deep	68	5.03	0.66

TABLE 2 (Continued)

	N	Mean	SD
Learn_Relational	68	5.02	0.74
GO_Relational	68	4.97	0.83
Assessment_Improvement	68	4.94	0.87
Purpose_ISE	67	4.81	0.83
Responsibility_Teaching	67	4.74	0.96
Curriculum_SocialReconstruct	67	4.73	0.90
Teach_SocialReform	68	4.71	0.87
Curriculum_Individualization	67	4.59	0.93
SE_Management	68	4.44	0.80
SE_instruction	68	4.28	0.69
Responsibility_Achievement	67	4.14	1.00
Responsibility_Motivation	67	4.04	1.07
Assessment_DemoLearning	68	3.90	0.86
Curriculum_Systematic	67	3.69	0.79
Assessment_Accountability	68	3.36	0.77
GO_AbilityAppr	68	3.29	1.01
Assessment_Hindrance	68	3.12	0.81
Learn_Surface	68	2.87	0.89
PUVandProfStatus	67	2.18	0.81
GO_WorkAvoid	68	2.00	0.71
InnateAbility	68	1.91	0.67
GO_AbilityAvoid	68	1.69	0.56
Epistemology_Simple	68	1.52	0.47
Effort	68	1.22	0.31

selected one class of their choosing to undertake the survey, within the period in which other measures were collected. Each participant was provided with a link to an online survey to be administered to their individual class. To preserve anonymity, teachers provided the link to their students, who used classroom computers during a regular teaching session to complete the survey online, with results sent directly to the researchers. The 35 items in the Tripod survey are organised around seven elements of effective classroom teaching, dubbed the seven Cs: care, control, clarify, challenge, captivate, confer and consolidate. Exploratory factor analysis of the survey items within this study's population did not replicate the presumed factor structure of the Tripod instrument, but instead identified a two-factor model with better fit, which included classroom management (control) and general instructional practices. The class average for each of the items within each of the two survey factors was averaged to provide an overall score for each teacher's student perceptions of classroom management and general instructional practices. Descriptive statistics and the factor structure for this measure are illustrated in Tables 4 and 5.

The two-factor model produced a strong fit, with estimates of reliability that were sufficiently high to have confidence in the total scores within each of the two identified factors (general instructional practices $\alpha = 0.97$, classroom management $\alpha = 0.82$).

TABLE 3 Higher-order factor structure and loadings.

	Social/ deep	Responsibility	Surface/ fixed	Efficacy/ intrinsic motivation	Pro-assessment
Curriculum—social reconstruction	0.95	−0.03	0.04	−0.23	0.06
Curriculum—individualization	0.79	−0.08	0.02	−0.13	−0.15
Teach—social reform	0.67	0.00	0.04	0.13	0.11
Curriculum—cross-curricular	0.65	−0.03	0.09	−0.01	0.01
Learning—deep	0.53	0.10	0.00	0.19	0.19
Goal orientation—relational	0.52	0.25	−0.09	−0.04	−0.15
Learning—relational	0.43	0.15	−0.10	0.26	0.31
Purpose—social utility value	0.39	−0.07	0.05	0.17	−0.10
Goal orientation—mastery approach	0.37	0.15	−0.19	0.39	−0.14
Responsibility for motivation	−0.04	0.91	0.12	−0.22	−0.01
Responsibility for achievement	−0.04	0.87	0.01	−0.02	−0.18
Responsibility for teaching	−0.08	0.81	−0.02	0.12	0.05
Responsibility relationships	0.30	0.72	0.05	−0.18	0.02
Teaching—deep	0.14	0.35	0.14	0.19	0.13
Learning—surface	0.03	0.05	0.67	0.25	0.10
Curriculum—systematic	0.23	−0.15	0.62	0.23	−0.09
Goal orientation—ability approach	0.22	0.09	0.55	−0.28	−0.16
Innate ability	−0.17	−0.02	0.46	−0.10	0.03
Goal orientation—ability avoidance	−0.06	0.18	0.41	−0.33	0.17
Epistemology—simple	−0.47	−0.01	0.40	0.10	0.12
Purpose—personal utility value and professional status	0.27	−0.23	0.38	0.14	0.10
Effort	−0.19	−0.06	0.35	0.03	0.02
Expectations (reverse coding)	−0.05	0.54	−0.34	0.23	0.13
Goal orientation—work avoidance	−0.08	0.05	0.32	−0.06	0.81
Self-efficacy—classroom management	−0.06	−0.08	0.07	0.86	−0.03
Self-efficacy—instruction	−0.10	0.10	0.04	0.79	−0.12
Purpose—intrinsic motivation and self-efficacy	0.13	−0.16	0.23	0.41	−0.12
Assessment—hindrance (reverse coded)	0.01	0.01	−0.09	−0.11	0.73
Assessment—demonstration of learning	−0.20	0.19	0.28	0.18	−0.41
Assessment—improvement	0.10	0.16	0.25	0.16	−0.28
Assessment—accountability for students and schools	−0.13	0.19	0.52	0.17	−0.05

Factor correlations

	Social/ deep	Responsibility	Surface/ fixed	Efficacy/ intrinsic motivation	Pro-assessment
Social/deep	1.00				
Responsibility	0.46	1.00			
Surface/fixed	−0.08	−0.09	1.00		
Efficacy/intrinsic motivation	0.30	0.16	0.17	1.00	
Pro-assessment	0.05	0.27	0.22	0.28	1.00

TABLE 4 Tripod descriptive statistics.

	N	Mean	SD
I like the ways we learn in this class	61	3.84	0.52
Students get to decide how activities are done in this class	61	2.92	0.53
My teacher wants us to share our thoughts	61	4.26	0.51
My teacher asks questions to be sure we are following along when s/he is teaching	61	4.26	0.41
My teacher asks students to explain more about the answers they give	61	4.11	0.40
My teacher in this class makes me feel that s/he cares about me	61	3.86	0.59
Our class stays busy and does not waste time	61	3.40	0.56
My teacher knows when the class understands, and when we do not	61	3.85	0.43
If I do not understand something, my teacher explains it another way	61	4.02	0.49
My teacher makes learning enjoyable	61	3.71	0.65
Students in this class treat the teacher with respect	61	3.79	0.58
My teacher does not let people give up when the work gets hard	61	4.06	0.45
My teacher wants me to explain my answers—why I think what I think	61	4.10	0.46
My teacher has several good ways to explain each topic that we cover in this class	61	3.83	0.45
The comments that I get on my work in this class help me understand how to improve	61	3.87	0.47
Student behaviour in this class is a problem	61	3.47	0.71
When s/he is teaching us, my teacher thinks we understand when we do not	61	3.49	0.54
We get helpful comments to let us know what we did wrong on assignments	61	3.94	0.47
Student behaviour in this class is under control	61	3.59	0.59
My teacher tries to understand how students feel about things	61	3.86	0.52
I hate the way that students behave in class	61	3.57	0.67
My teacher makes lessons interesting	61	3.65	0.57
My classmates behave the way my teacher wants them to	61	3.48	0.56
This class does not keep my attention—I get bored	61	3.36	0.58
My teacher takes the time to summarise what we learn each day	61	3.66	0.53
In this class, we learn a lot almost every day	61	3.80	0.50
My teacher seems to know if something is bothering me	61	3.23	0.50
My teacher checks to make sure we understand when s/he is talking	61	3.98	0.42
My teacher respects my ideas and suggestions	61	4.06	0.52
Student behaviour in this class makes the teacher angry	61	3.46	0.68
My teacher explains difficult things clearly	61	3.85	0.45
In this class, my teacher accepts nothing less than our full effort	61	3.94	0.41
My teacher gives us time to explain our ideas	61	3.99	0.47
In this class, we learn to correct our mistakes	61	3.94	0.42
Students speak up and share their ideas about classwork	61	3.80	0.51

Teaching practice

Teaching practice was measured using the Visible Classroom program to identify the use of 16 specific teaching practices supported by meta-analytic research as promoting higher levels of academic achievement (Hattie, 2009; Skipp & Tanner, 2015). Participants in the study audio-recorded themselves delivering between three and five full lessons with the

TABLE 5 Revised Tripod survey factor structure and loadings.

	Original Tripod survey classification	General instructional practices factor	Classroom management factor
My teacher really tries to understand how students feel about things	Care	0.85	−0.03
My teacher explains difficult things clearly	Clarify	0.83	−0.00
My teacher checks to make sure we understand when s/he is talking	Consolidate	0.83	−0.03
My teacher gives us time to explain our ideas	Confer	0.82	−0.02
My teacher has several good ways to explain each topic that we cover in this class	Clarify	0.81	−0.01
If I do not understand something, my teacher explains it another way	Clarify	0.80	0.01
My teacher respects my ideas and suggestions	Confer	0.80	0.01
In this class, we learn to correct our mistakes	Consolidate	0.79	0.02
My teacher makes lessons interesting	Captivate	0.79	0.00
My teacher knows when the class understands, and when we do not	Clarify	0.79	−0.06
My teacher does not let people give up when the work gets hard	Challenge	0.78	−0.00
I like the ways we learn in this class	Captivate	0.78	0.00
My teacher makes learning enjoyable	Captivate	0.77	0.05
In this class, we learn a lot almost every day	Challenge	0.77	0.03
My teacher in this class makes me feel that s/he really cares about me	Care	0.77	0.02
My teacher wants me to explain my answers—why I think what I think	Challenge	0.76	−0.02
We get helpful comments to let us know what we did wrong on assignments	Consolidate	0.73	0.04
My teacher asks questions to be sure we are following along when s/he is teaching	Challenge	0.73	−0.02
The comments that I get on my work in this class help me understand how to improve	Consolidate	0.71	0.04
My teacher seems to know if something is bothering me	Care	0.71	−0.10
In this class, my teacher accepts nothing less than our full effort	Challenge	0.70	−0.03
My teacher asks students to explain more about the answers they give	Challenge	0.70	−0.06
My teacher takes the time to summarise what we learn each day	Consolidate	0.68	−0.04
My teacher wants us to share our thoughts	Confer	0.66	0.04
Students speak up and share their ideas about class work	Confer	0.60	0.08
Students get to decide how activities are done in this class	Confer	0.47	0.07
This class does not keep my attention—I get bored	Captivate	−0.28	−0.27
Student behaviour in this class is a problem	Control	0.22	−0.87
I hate the way that students behave in class	Control	0.24	−0.77
Student behaviour in this class makes the teacher angry	Control	0.09	−0.71

TABLE 5 (Continued)

	Original Tripod survey classification	General instructional practices factor	Classroom management factor
My classmates behave the way my teacher wants them to	Control	0.28	0.52
Student behaviour in this class is under control	Control	0.24	0.45
Students in this class treat the teacher with respect	Control	0.33	0.45
Our class stays busy and does not waste time	Control	0.35	0.42
When s/he is teaching us, my teacher thinks we understand when we do not	Clarify	−0.07	−0.41

same class. Participants were able to select the lessons they wished to record and were not required to record the lessons sequentially. They were encouraged to limit the time between recordings to no more than three lessons. Recordings were then uploaded through a Visible Classroom mobile phone-based app, which auto-transcribed and anonymised the lessons before they were shared with coders. Lesson transcripts were then coded by trained assessors from the University of Melbourne's Centre for Program Evaluation, which operated Visible Classroom and also provided training for these assessors. Frequency counts for each identified practice were averaged across the lessons.

The descriptive statistics and factor structure for this measure are illustrated in Tables 6 and 7. Exploratory factor analysis by Witter (2023b) identified a higher-order structure of two broad instructional elements: the first of these elements related to teaching practices that focused on directing and correcting students in the classroom, including providing directions, offering feedback and correcting inappropriate behaviour; the second of these elements related to instructional teaching techniques such as introducing and explaining content, convergent and divergent questioning, summarising and reviewing key ideas and reviewing content. This model excluded two of the original 16 teaching practices (goals/success criteria and connections) because neither scale contributed to either factor. The correlation between the two factors was $r = 0.27$. Estimates of reliability were sufficiently high to support these two factors (explicit instruction $\alpha = 0.71$, directions and corrections $\alpha = 0.69$).

Analytical approach

Hierarchical cluster modelling (HCM) was utilised to determine whether particular profiles of teacher types could be identified based on their results across the aforementioned measures. HCM was used with between-groups linkage as well as *k*-means cluster analysis using Ward's method to identify several clusters with sufficiently distinct characteristics to highlight multiple distinct profiles of teachers based on their beliefs and practices, and their students' perceptions of teaching. This method sought to minimise the distance between subjects within the cluster (by reducing the variance within the group) and avoid 'long chaining' (Aldenderfer & Blashfield, 1984). The Euclidean distance was used as a similarity measure across the five higher-order teacher beliefs factors, the two higher-order Visible Classroom teaching factors and the two higher-order student perceptions of teaching factors. Cluster analysis only included participants who had provided data for all measures, which reduced the overall sample to 58 participants. Statistical analyses were conducted using the Statistical Package for the Social Sciences (SPSS).

TABLE 6 Practice overview and descriptive statistics for Visible Classroom practices.

Practice	Overview of practice	*N*	Mean	SD
Repeats comment	Repeats comment or question from student before answering to encourage and reinforce correct answers	59	25.52	10.65
Convergent questioning	Checks fact recall by asking closed questions (characterised by YES/NO questions/quizzes), or questions that have one correct answer or a designated set of possible answers	59	17.72	9.36
Prompting	The teacher uses prompting or probing (a question or statement used to elicit an appropriate student response)	59	5.25	3.25
Positive environment	Creates a classroom environment in which students have an opportunity to freely/spontaneously ask/provide task-related questions/feedback that is content specific	59	5.50	2.35
Deepen understanding	Provides students with the opportunity to deepen understanding and make connections between ideas and build on prior knowledge through sustained context-specific dialogue and open-ended or divergent questions, to which there is no one correct answer	59	6.14	3.56
Introduces explains	Introduces and explains new/complicated vocabulary/terminology, simplifies concepts by breaking into different levels or elaborates on abstract concepts using concrete, developmentally appropriate and high-quality examples	59	10.02	4.63
Important information	Emphasises important points	59	2.38	1.57
Summarise	Concludes the lesson by recapitulating/summarising key points AND provides the opportunity for follow-up/future engagement	59	1.39	1.07
Review	Reviews previously learned concepts at the beginning of and/or throughout the lesson or builds upon prior/assumed knowledge	59	2.06	1.14
Instructions	Provides step-by-step instructions on completing tasks/activities	59	11.88	4.90
Behaviour	Sets clear behaviour expectations and prompts behavioural reminders during the lesson	59	3.23	2.32
Resources	Reviews previously learned concepts at the beginning of and/or throughout the lesson, or builds upon prior/assumed knowledge	59	4.16	1.50
Scaffolded activities and collaboration	Provides opportunities for collaborative, scaffolded activities with their peers to actively engage in learning	59	2.01	1.61
Feedback	The teacher provides immediate, specific and corrective feedback to individual students/groups	59	4.71	2.34

Note: Numbers provided are based on an individual teacher average of identified practices per lesson. All participants taught between three and five lessons in total.

RESULTS

As a precursor to HCM, MANOVA (multivariate analysis of variance) was conducted to validate the approach this study took to ensure that subject area differences did not significantly moderate outcomes across higher-order factors. Using the clustering variables planned for HCM as dependent variables, the fixed factor of teachers' general teaching areas was tested. All η^2 values were small and not statistically significant, suggesting no

TABLE 7 Factor structure and loadings of Visible Classroom practices.

Practice	Explicit instruction	Directions and corrections
Repeats comment	0.99	−0.11
Convergent questioning	0.68	0.12
Prompting	0.62	0.04
Positive environment	0.51	0.12
Deepen understanding	0.50	−0.13
Introduces explains	0.33	−0.06
Important information	0.27	−0.20
Summarise	0.21	−0.30
Review	0.13	0.06
Instructions	−0.03	0.88
Behaviour	0.16	0.65
Resources	−0.15	0.64
Scaffolded activities and collaboration	0.07	0.58
Feedback	0.23	0.28

interactions between subject areas and teachers' beliefs, their use of Visible Classroom practices or their students' perceptions of teaching. These results should be interpreted with caution, however, given the sample size of each subgroup by teaching area, but were sufficient to proceed with the cluster analysis without segmentation based on teaching area.

This study used a heuristic approach to determining the optimal number of clusters. The ultimate goal of cluster validation in this study was to create a balance between identifying statistical indices that indicated the highest density within clusters and distances between clusters, with meaningful distinctions between clusters based on the higher-order factors examined within the study. Sample size was a potential impediment to identifying a larger number of clusters, and may have steered statistical indices towards the smallest number of clusters for the model. This was suggested by the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), which were the lowest in two-cluster and three-cluster models, respectively. Additionally, tests of the silhouette measure for cohesion and separation suggested a limited distinction between models of two to five clusters, all of which were fair. Analysis of variance (ANOVA) conducted as part of the *k*-means test of the two-cluster model found significant interactions at the $p = 0.05$ level for three of nine factors, while the three-cluster model found significant interactions between four factors at the $p = 0.05$ level and one factor at the $p = 0.10$ level. The two-cluster model split into groups of 27 and 31 teachers, while the three-cluster model split into groups of 20, 28 and 10 teachers. Across both the two-cluster and three-cluster models, the 31-teacher cluster remained mostly stable, with all but four teachers in the same cluster in both models. In the three-cluster model, the majority of members in the additional cluster came from a splintering of the 28-teacher cluster from the two-cluster model. A four-cluster model was additionally tested but demonstrated progressively poorer statistical indices than the other models (as previously noted), and splintered into a fourth cluster size that was deemed too small ($n = 2$) to meaningfully represent an important subset of the overall population. For this reason, the three-cluster model was deemed most appropriate for the study. Table 8 provides descriptive statistics for the three-cluster model, while Table 9 describes ANOVA results.

TABLE 8 Higher-order factor descriptive statistics for three clusters.

Factors	Cluster	Mean	SD	*N*
Visible Classroom—explicit instruction	1	11.08	2.01	20
	2	6.05	1.33	28
	3	10.06	2.12	10
Visible Classroom—directing and correcting	1	4.66	1.01	20
	2	4.70	1.41	28
	3	8.07	1.57	10
Efficacy/intrinsic motivation	1	4.67	0.49	20
	2	4.52	0.65	28
	3	4.16	0.62	10
Social/deep	1	4.97	0.52	20
	2	5.13	0.46	28
	3	4.91	0.56	10
Surface/fixed	1	2.20	0.42	20
	2	2.32	0.33	28
	3	2.01	0.25	10
Pro-assessment	1	3.89	0.64	20
	2	4.21	0.52	28
	3	3.89	0.58	10
Responsibility	1	4.52	0.65	20
	2	4.58	0.78	28
	3	4.24	1.09	10
Tripod—general instructional	1	4.00	0.35	20
	2	3.77	0.45	28
	3	3.58	0.34	10
Tripod—class control	1	3.98	0.36	20
	2	3.72	0.45	28
	3	3.49	0.30	10

TABLE 9 ANOVA results using cluster type as the criterion.

	Sum of squares	df	Mean square	*F*	*p*
Visible Classroom—explicit instruction	162.70	2	3.00	54.26	<0.01
Visible Classroom—directing and correcting	47.39	2	1.73	27.32	<0.01
Efficacy/intrinsic motivation	0.88	2	0.35	2.49	0.09
Social/deep	0.24	2	0.25	0.96	0.39
Surface/fixed	0.38	2	0.12	3.07	0.05
Pro-assessment	0.76	2	0.33	2.28	0.11
Responsibility	0.44	2	0.63	0.70	0.50
Tripod—general instructional	0.87	2	0.16	5.48	<0.01
Tripod—class control	2.61	2	0.25	10.58	<0.01

The major dimensions discriminating among the three clusters were the frequency of the use of general instructional practices and directing and correcting practices (as measured by Visible Classroom), students' perceptions of both general instructional practices and classroom management and teachers' efficacy/intrinsic motivational beliefs, surface/fixed beliefs and pro-assessment beliefs, although the latter two lacked statistical power, potentially due to the numbers of teachers within clusters.

MANOVA produced similar results to ANOVA in terms of significant interactions between cluster type and higher-order factors, and included post-hoc analyses, which indicated several significant differences between particular pairs of clusters, in addition to numerous non-significant but still moderate to large differences in clusters. Table 10 identifies significant interactions between cluster type and higher-order factors, as well as Cohen's *D*, to illustrate the magnitude of differences between each cluster group against each of the factors examined in this study.

The first cluster of teachers illustrated a pattern of greater frequency of explicit instructional practices and directing and correcting strategies, accompanied by significantly higher perceptions of general instructional quality and class control, with medium differences between Clusters 1 and 2 in both student perception of general instruction and class control, and large differences between Clusters 1 and 3, as measured by Cohen's *D*. Cluster 1 had the highest self-efficacy and personal responsibility overall but held no beliefs that significantly distinguished them from the other two groups, and held few beliefs that were moderately or largely different (based on Cohen's *D*) from either of the groups, with the exception of large-sized differences in self-efficacy beliefs compared to Cluster 1 and medium-sized lower pro-assessment beliefs than Cluster 2.

The second cluster of teachers was distinguished by a few key characteristics. First, they engaged in relatively less frequent use of both types of Visible Classroom practices compared to Clusters 1 and 3, which amounted to both large and statistically significant differences. A notable finding from Cluster 2 was that in contrast to Cluster 1, which had the highest use of Visible Classroom practices and the highest levels of student perception feedback, Cluster 2 outperformed Cluster 3 in terms of student feedback, whilst simultaneously outperforming Cluster 3 on both Tripod factors. This cluster, like Cluster 1, also held relatively high self-efficacy/intrinsic motivation beliefs compared to Cluster 3, and the highest overall social/deep, pro-assessment beliefs and surface/fixed beliefs, although only the latter was statistically significantly different from either cluster.

The third cluster was distinguished by their relatively high use of both types of Visible Classroom practices, particularly directing and correcting strategies, which they used far more frequently than either cluster. This group also had comparatively lower levels of self-efficacy/intrinsic motivation, lower surface/fixed beliefs and a lower sense of personal responsibility for teaching compared to the other two groups, although these beliefs differences were statistically non-significant. Although this cluster analysis did not seek out causal relationships between factors, its results were consistent with additional research by Witter (2023b) suggesting that higher frequencies of directing and corrective practices correspond to lower levels of student perceptions of teacher quality, and also supported research indicating that higher levels of self-efficacy predict higher levels of quality teaching and learning (Klassen & Tze, 2014).

DISCUSSION

Efforts to synthesise the profiles of effective teachers, their qualities, behaviours and outcomes, are a kind of Holy Grail in teacher quality research. While this study cannot lay claim to uncovering a clear and indefensible profile of teacher quality, the results do

TABLE 10 Multiple comparisons.

Factors	Cluster		Mean difference	Sig.	*D*
Visible Classroom—explicit instruction	1	2	5.03	0.00	2.95
		3	1.02	0.29	0.49
	2	1	−5.03	0.00	−0.2.95
		3	−4.01	0.00	−0.2.26
	3	1	−1.02	0.29	−0.49
		2	4.01	0.00	2.26
Visible Classroom—directing and correcting	1	2	−0.04	0.99	−0.03
		3	−3.41	0.00	−2.58
	2	1	0.04	0.99	0.03
		3	−3.37	0.00	−0.2.26
	3	1	3.41	0.00	2.58
		2	3.37	0.00	2.26
Efficacy/intrinsic motivation	1	2	0.16	0.64	0.27
		3	0.51	0.08	0.92
	2	1	−0.16	0.64	−0.27
		3	0.36	0.24	0.56
	3	1	−0.51	0.08	−0.92
		2	−0.36	0.24	−0.56
Social/deep	1	2	−0.33	0.52	−0.33
		3	0.05	0.96	0.10
	2	1	0.16	0.52	0.33
		3	0.21	0.48	0.41
	3	1	−0.05	0.96	−0.10
		2	−0.21	0.48	−0.41
Surface/fixed	1	2	−0.13	0.45	−0.33
		3	0.19	0.35	0.55
	2	1	0.13	0.45	0.33
		3	0.32	0.05	1.11
	3	1	−0.19	0.35	−0.55
		2	−0.32	0.05	−1.11
Pro-assessment	1	2	−0.32	0.15	−0.55
		3	0.00	1.00	0.01
	2	1	0.32	0.15	0.55
		3	0.33	0.28	0.59
	3	1	0.00	1.00	−0.01
		2	−0.33	0.28	−0.59
Responsibility	1	2	−0.06	0.97	−0.08
		3	0.29	0.62	0.32
	2	1	0.06	0.97	0.08
		3	0.34	0.47	0.36
	3	1	−0.29	0.62	−0.32
		2	−0.34	0.47	−0.36

TABLE 10 (Continued)

Factors	Cluster		Mean difference	Sig.	*D*
Tripod—general instructional	1	2	0.23	0.13	0.57
		3	0.42	0.03	1.21
	2	1	−0.23	0.13	−0.57
		3	0.19	0.42	0.47
	3	1	−0.42	0.03	−1.21
		2	−0.19	0.42	-.0.47
Tripod—class control	1	2	0.27	0.07	0.65
		3	0.49	0.01	1.47
	2	1	−0.27	0.07	−0.65
		3	0.22	0.29	0.58
	3	1	−0.49	0.01	−1.47
		2	−0.22	0.29	−0.58

suggest a particular profile of teacher quality that showcases the interplay between a set of critical beliefs, behaviours and student outcomes. An emergent profile this research identified (from Cluster 1) is that of an active teacher who leverages high-frequency explicit instructional practices, holds higher levels of self-efficacy and intrinsic motivation for teaching and personal responsibility and produces higher perceptions of teacher quality across multiple aspects of classroom practice. The practices of these teachers (and relative frequency of their use) and other characteristics, including their students' perceptions of teaching quality, were consistent with research that promotes conceptions of quality teaching as 'active' versus 'passive' (Brophy, 1986). Particular active teaching practices, such as those measured as Visible Classroom explicit instructional practices, are consistently among the highest-impact classroom-based influences on student achievement (Hattie, 2009; Muijs, 2006). Central to this theory is the notion that whether leading whole-group instruction or supporting individuals or small groups of learners, effective teachers are constantly at work throughout their lessons, making numerous and frequent teaching moves to ensure that their classrooms move at a brisk pace and that time on task is maximised so that students spend as much time as possible engaged in learning experiences.

At the same time, Visible Classroom practices characterised as 'directing and correcting', which were selected on the basis of their empirical support as also promoting student achievement, did not emerge within this analysis as a distinguishing factor between Clusters 1 and 2, and the fact that the cluster with the poorest student perception feedback (Cluster 3) was also the one whose teachers used these types of practices most frequently raises questions regarding whether or not these measures should be accepted as quality teaching practices or not. Cluster 3 provided a further contrasting profile from both Clusters 1 and 2, as teachers in this group were also lower in their self-efficacy and intrinsic motivation, their sense of personal responsibility and students' perceptions of teaching quality. While the first and second clusters of teachers had limited differences in terms of their beliefs, they were distinguished by their practices (comparatively less frequent use of explicit instructional practices) and by their students' relatively lower (compared to Cluster 1, but not Cluster 3) perceptions of teaching quality.

This study sought to situate teacher beliefs research within a teacher quality paradigm, and to address the lack of empirical research on teacher beliefs—and particularly teacher belief systems—within studies of teacher quality. Each cluster in this study arguably represented

a type of beliefs system, a complex, integrated collection of intersecting beliefs. The study illustrated how those beliefs systems integrate and connect with other teacher quality factors within a given cluster of teachers. Cluster analysis illustrated how the strength of some beliefs within these systems tended to distinguish teachers more than others, and also how—across particular clusters—beliefs around efficacy, intrinsic motivation and personal responsibility accompanied differential practices and students' perceptions of teaching.

Teacher beliefs are an important facet of teacher quality and play an important role in shaping practices that promote optimal learning outcomes (Fives & Buehl, 2012), yet beliefs are inadequate to look at in isolation when attempting to gain insights into what makes a quality teacher. As Goe (2007) noted, teacher quality encompasses inputs such as beliefs, processes that reflect quality teaching practices and positive student outcomes. This study's results illustrate the importance of drawing upon inputs (beliefs), processes (quality teaching practices) and outcomes (student perceptions of quality) in an effort to develop a profile of teacher quality, while the use of cluster analysis—an underutilised methodology in teacher quality research—has proven itself to be a valuable means for finding commonalities between different groups of teachers in a way that reflects the complex nature of the teacher's role.

LIMITATIONS AND FUTURE RESEARCH

The relatively low sample size of this study may have limited the number of unique profiles that could be identified as clusters, and as such it is possible that other profiles were not captured in the analysis, or that particular profiles were shaped around small and idiosyncratic clusters of teachers. It may also have prevented sufficient detection of possible interactions between factors such as teaching areas. Though there is no clear consensus regarding how best to establish statistical power in cluster analysis and what thresholds for cluster size should be viewed as minimal or optimal, Dalmaijer et al. (2022) suggested that clusters of $N = 20$ to $N = 30$ can generally yield sufficient statistical power. In this study, two of three clusters fell within those guidelines and one did not. Future attempts to apply cluster analysis methods to teacher quality factors would benefit from at least somewhat larger sample sizes. Larger sample sizes would also likely yield improved statistical indices when attempting to identify larger numbers of distinct clusters and increase the likelihood of detecting statistical effects. Given the magnitude of differences that were identified between the three clusters against the nine higher-order factors, the majority of which were medium to large, larger sample sizes would increase confidence that these differences were statistically significant as well. Nevertheless, although the sample size likely reduced statistical power in this study, it was adequate to provide sufficient power to detect many statistical effects across beliefs, practices and students' perceptions of teaching.

While this study sought to expand upon previous papers that have targeted a narrower range of teacher quality components, it is still limited by what factors it has included and excluded. Although a wide range of teacher beliefs factors was tested within the study, the field has expanded to include far too many possible constructs to be examined within a single study (and counting), and thus future research should attempt to incorporate contemporary and promising new constructs that present themselves as strong potential influences on teacher quality. Yet, at the same time, the beliefs structures examined within this study provided a unique and rare integration of teacher beliefs systems measurement and analysis within teacher quality research. This study did not provide an incontrovertible profile of teachers' belief systems, but illustrated how cluster analysis can be used to deepen insights into how teachers integrate numerous complex and even potentially incongruent beliefs.

Given the limited empirical studies undertaken in this area, far more research must be undertaken to examine how teacher belief systems interact with facets of teacher quality. Although this study produced a potential profile of a higher-quality teacher based on some of the factors it investigated, this and other models incorporating features not examined in this study should be a focus of future studies.

The use of Visible Classroom as a mechanism for measuring teacher practice was valuable in that it offered a low-inference means of measuring quality teaching, focusing on frequency—but at the same time, such measures are limited in that they emphasise 'how much' versus 'how well' teachers employ particular practices in the classroom. A logical next step would be to incorporate higher-inference observational measures of teaching practice that have already been empirically linked to other key measures of quality teaching, such as student academic gains. Teachers' value-added to academic achievement is a frequent and arguably essential component of teacher quality measurement, and its absence as a measure in this study was a clear limitation. While certain factors and measures within this study were selected because of extant research indicating that they are predictive of higher levels of student achievement, these relationships still tend to be weak to moderate at best (Kane & Staiger, 2012), and an evidence base working to establish a teacher quality profile will be far richer when it includes direct measures of student achievement.

CONCLUSION

This study produced a profile of teacher quality drawn from a wide range of important beliefs, practices and student perceptions of teaching, contemporising the extant research profiles of quality teachers that are primarily derived from literature reviews by examining a range of characteristics within a cluster analysis approach, painting a profile of quality teachers that encompasses key sets of beliefs, particular behaviours (and their frequency) and students' perspectives on their teaching quality. These teachers may share many of the same beliefs as teachers who behave and are perceived quite differently, but they stand out in terms of their self-efficacy and intrinsic motivation to teach.

The results of this study illustrate how active teaching resonates with students, but also raise questions regarding what types of active teaching practices best support students' perceptions of effective teaching. Given the complexities surrounding the roles of teachers and the unique contextual factors that shape each classroom, it would be unreasonable to conclude that there is just one singular profile of quality teaching for all students everywhere. But we must balance that acknowledgement with equal recognition that there are elements of teacher quality that do appear to resonate across many such contexts, which was a key aim of this study. Its findings indicate that certain aspects of teacher quality are relevant, applicable and appropriate across contexts, and illustrated that they may work together to reflect a more holistic understanding of the quality teacher and their various dimensions.

FUNDING INFORMATION
Not applicable.

CONFLICT OF INTEREST STATEMENT
No potential conflict of interest was reported by the authors.

DATA AVAILABILITY STATEMENT
The data supporting the findings of this study are available within the paper, with additional supplementary materials provided at OSF via the following link: https://osf.io/28zuf/.

Raw data for this study were generated at the Melbourne Graduate School of Education. Additional derived data supporting the findings of this study are available on request from Dr Michael Witter.

ETHICS STATEMENT

The study presented within this paper was conducted in accordance with the *British Educational Research Journal*'s Ethical Guidelines for Educational Research. This research was designed to be inclusive of a range of stakeholders, with emphasis on respecting and prioritising the privacy and welfare of each participant. The research design explicitly sought to maximise benefit to both teachers and students involved in the study, and to minimise harm. This study was developed in alignment with the Australian Code for the Responsible Conduct of Research. The study went through three separate ethics approval processes prior to its implementation. Initially, it was approved by the Melbourne Graduate School of Education's Human Ethics Advisory Group, with revisions to ensure all ethical requirements were fulfilled, with additional revisions and approval provided via Melbourne University's Humanities and Applied Sciences Human Ethics Sub-Committee. Further approval was obtained for each state or territory in which research was conducted via the ACT Education Directorate, the Victorian Department of Education and Training, the Western Australia Department of Education and the Northern Territory Department of Education.

ORCID

Michael Witter https://orcid.org/0000-0002-7606-3070

REFERENCES

Aldenderfer, M., & Blashfield, R. (1984). *Cluster analysis*. Sage.

Breault, R. A. (2013). "She was great, but…": Examining preservice recollections of favorite and most effective teachers. *Professional Educator*, *37*(1), art. 1.

Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, *41*(10), 1069–1077.

Brophy, J. E., & Good, T. L. (1984). *Teacher behavior and student achievement*. Institute for Research on Teaching, Michigan State University.

Cochran-Smith, M. (2003). The unforgiving complexity of teaching: Avoiding simplicity in the age of accountability. *Journal of Teacher Education*, *54*(1), 3–5.

Colker, L. J. (2008). Twelve characteristics of effective early childhood teachers. *YC Young Children*, *63*(2), 68–73.

Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2022). Statistical power for cluster analysis. *BMC Bioinformatics*, *23*(1), 1–28.

Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, *8*(1), 1–44.

Ferguson, R. F., & Danielson, C. (2015). How framework for teaching and tripod 7Cs evidence distinguish key components of effective teaching. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the measures of effective teaching project* (pp. 98–143). Wiley.

Fives, H., & Buehl, M. M. (2012). *Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us?* American Psychological Association.

Gay, M. J. (2012). *Excellent teaching: A collective case study of outstanding elementary mathematics teachers' teaching of mathematics*. The University of Nebraska-Lincoln.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. National Comprehensive Center for Teacher Quality.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. National Comprehensive Center for Teacher Quality.

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.

Hanushek, E. (2002). Teacher quality. In L. T. Izumi & W. M. Evers (Eds.), *Teacher quality* (pp. 1–12). Hoover Press.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Research Paper, MET Project, Bill & Melinda Gates Foundation.

Keeley, J. W., Ismail, E., & Buskist, W. (2016). Excellent teachers' perspectives on excellent teaching. *Teaching of Psychology*, *43*(3), 175–179.
Kennedy, M. M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher*, *39*(8), 591–598.
Klassen, R. M., & Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76.
LaFee, S. (2014). Students evaluating teachers. *The Education Digest*, *80*(3), 4.
Muijs, D. (2006). Measuring teacher effectiveness: Some methodological reflections. *Educational Research and Evaluation*, *12*(1), 53–74.
Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Economic Policy Institute.
Skipp, A., & Tanner, E. (2015). *The visible classroom: Evaluation report and executive summary*. Education Endowment Foundation.
Stronge, J. H. (2018). *Qualities of effective teachers*. ASCD.
Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, *62*(4), 339–355.
Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122.
Witter, M. S. (2023a). *Teacher belief systems: An empirical investigation* (article in preparation).
Witter, M. S. (2023b). *Teacher belief systems and teacher quality: A multidimensional analysis* (article in preparation).

How to cite this article: Witter, M. & Hattie, J. (2024). Can teacher quality be profiled? A cluster analysis of teachers' beliefs, practices and students' perceptions of effectiveness. *British Educational Research Journal*, *50*, 653–675. https://doi.org/10.1002/berj.3938